

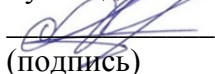


МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Дальневосточный федеральный университет»  
(ДВФУ)

**ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ (ШКОЛА)**

СОГЛАСОВАНО

Руководитель ОП

 Чеботарев А.Ю.  
(подпись) (ФИО)

«УТВЕРЖДАЮ»

И.о. директора департамента  
математического и компьютерного  
моделирования  
 Сущенко А.А.  
(подпись) (ФИО)

«27» сентября 2021 г.



**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Компьютерные методы анализа больших данных

**Направление подготовки 01.04.02 Прикладная математика и информатика**  
(Математическое моделирование (совместно с ИПМ ДВО РАН))

**Форма подготовки очная**

курс 1 семестр 1

лекции 16 час.

практические занятия не предусмотрены

лабораторные работы 34 час.

в том числе с использованием МАО лек. 16 / пр. - / лаб. 34 час.

всего часов аудиторной нагрузки 50 час.

в том числе с использованием МАО 50 час.

самостоятельная работа 94 час.

в том числе на подготовку к экзамену 36 час.

контрольные работы (количество) не предусмотрены

курсовая работа / курсовой проект 1 семестр

зачет не предусмотрен

экзамен 1 семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки **01.04.02 Прикладная математика и информатика** утвержденного приказом Министерства образования и науки РФ от 10 января 2018 г. №13 (с изменениями и дополнениями)

Рабочая программа обсуждена на заседании департамента математического и компьютерного моделирования протокол № 1 от «27» сентября 2021 г.

И.о. директора департамента математического и компьютерного моделирования А.А. Сущенко  
Составитель (ли): Т.В. Пак

Владивосток

2021

**Оборотная сторона титульного листа РПД**

**I. Рабочая программа пересмотрена на заседании департамента:**

Протокол от «\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**II. Рабочая программа пересмотрена на заседании департамента:**

Протокол от «\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

## I. Цели и задачи освоения дисциплины:

Цель: Ознакомление с основными технологиями решения задач обработки больших данных и приобретение навыков в разработке приложений для аналитики больших данных.

Задачи:

- Ознакомить студентов с основными технологиями решения задач обработки больших по объему, быстро изменяющихся и плохо структурированных данных;
- Научить применять методы анализа больших данных;
- Дать навыки реализации приложения для аналитики больших данных.

Для успешного изучения дисциплины «Компьютерные методы анализа больших данных» у обучающихся должны быть сформированы следующие предварительные компетенции:

- ПК-1 Способен проводить научные исследования и получать новые научные и прикладные результаты самостоятельно и в составе научного коллектива

- ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности

Профессиональные компетенции выпускников и индикаторы их достижения:

Тип задач	Код и наименование профессиональной компетенции (результат освоения)	Код и наименование индикатора достижения компетенции
проектный	ПК-3 Способен управлять проектами, планировать научно-исследовательскую деятельность, анализировать риски, управлять командой проекта	ПК-3.1 Демонстрирует знание методов составления и контроля плана выполняемой научно-исследовательской работы, основы бизнес-планирования
		ПК-3.2 Использует методы математического моделирования, принятия решений, разбиения задачи на подзадачи, оценивает результат работы команды проекта, риски проекта, составляет бизнес-план
		ПК-3.3 Применяет методы математического моделирования, планирования научно-исследовательской деятельности, работы в научно-исследовательском коллективе

Код и наименование индикатора достижения компетенции	Наименование показателя оценивания (результата обучения по дисциплине)
ПК-3.1 Демонстрирует знание методов составления и контроля	Знает основные методы планирования и контроля НИР в области информационных технологий, основы бизнес-

Код и наименование индикатора достижения компетенции	Наименование показателя оценивания (результата обучения по дисциплине)
плана выполняемой научно-исследовательской работы, основы бизнес-планирования	планирования
	Умеет строить математические алгоритмы, модели и реализовывать их с помощью языков программирования
	Владеет навыками компьютерной обработки вычислительных задач
ПК-3.2 Использует методы математического моделирования, принятия решений, разбиения задачи на подзадачи, оценивает результат работы команды проекта, риски проекта, составляет бизнес-план	Знает профессиональную терминологию, содержание ключевых понятий и определений, используемых в теории и практике применения информационных технологий в науке и образовании
	Умеет применять математический язык, методы при построении моделей объектов профессиональной деятельности с использованием инструментальных средств компьютерного моделирования
	Владеет навыками использования прикладного программного обеспечения для решения задач в профессиональной деятельности
ПК-3.3 Применяет методы математического моделирования, планирования научно-исследовательской деятельности, работы в научно-исследовательском коллективе	Знает информационные ресурсы и базы данных по научно-исследовательской теме
	Умеет самостоятельно расширять и углублять знания в области информационных технологий
	Владеет навыками работы с программными продуктами и информационными ресурсами

## II. Трудоемкость дисциплины и видов учебных занятий по дисциплине

Общая трудоемкость дисциплины составляет 4 зачётные единицы 144 академических часа).

(1 зачетная единица соответствует 36 академическим часам)

Видами учебных занятий и работы обучающегося по дисциплине являются:

Обозначение	Виды учебных занятий и работы обучающегося
Лек	Лекции
Лаб	Лабораторные работы
СР	Самостоятельная работа обучающегося в период теоретического обучения
Контроль	Самостоятельная работа обучающегося и контактная работа обучающегося с преподавателем в период промежуточной аттестации

Структура дисциплины:

Форма обучения – очная.

№	Наименование раздела дисциплины	ме	ст	Количество часов по видам учебных занятий и работы обучающегося	Формы промежуточной
---	---------------------------------	----	----	---	---------------------

			Лек	Лаб	Пр	ОК	СР	Контроль	аттестации
1	Раздел 1. Парадигма MapReduce	1	8	22	-	-	58	36	УО-1; ПР-6; ПР-9
2	Раздел 2. Кластеризация больших данных	1	8	12	-	-	58	36	
	Итого:		16	36	-	-	58	36	

### III. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

**Лекционные занятия (16 час. в том числе 16 часов с использованием интерактивных методов обучения)**

#### **Раздел 1. Парадигма MapReduce (8 часов).**

**Тема 1. Распределенные файловые системы. Основы MapReduce. (2 часа)** Физическая организация вычислительных узлов. Организация файловой системы. Map-задачи. Reduce-задачи. Сумматоры. Детали исполнения MapReduce.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 2. Исполнение алгоритмов с использованием MapReduce. (2 часа)** Матрично-векторное умножение. Операции реляционной алгебры. Выбор, проекция, объединение, пересечение. Группировка и агрегация.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 3. Расширения к MapReduce. Модель стоимости связи. (2 часа)** Рекурсивные расширения. Pregel. Стоимость связи для сетей задач.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 4. Теория сложности для MapReduce. (2 часа)** Размер Reducer'a и скорость репликации. Графовые модели MapReduce задач. Схемы маппинга.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

#### **Раздел 2. Кластеризация больших данных (8 часов).**

**Тема 1. Введение в кластеризацию. Иерархические алгоритмы. (2 часа)** Точки, пространства и расстояния. Стратегии кластеризации. Проблема размерности. Иерархические алгоритмы в евклидовых и неевклидовых пространствах. Эффективность.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 2. Алгоритмы k-среднего. (2 часа)** Основы алгоритмов k-среднего. Инициализация кластеров для алгоритмов k-среднего. Выбор правильного значения k. Алгоритм Брэдли, Файяда и Рейны.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 3. Алгоритм CURE. Кластеризация в неевклидовых пространствах. (2 часа)** Инициализация и исполнение алгоритма CURE. Представление кластеров в алгоритме GRGPF. Инициализация дерева кластеров. Добавление точек в алгоритме GRGPF. Разделение и слияние кластеров.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

**Тема 4. Кластеризация для потоков и параллелизм. (2 часа)** Модель потокового вычисления. Алгоритм кластеризации потока. Кластеризация в параллельной среде.

Занятие проводится с использованием метода интерактивного обучения «лекция-консультация».

#### **IV. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА И САМОСТОЯТЕЛЬНОЙ РАБОТЫ**

**Лабораторные работы (34 часов в том числе 34 часа с использованием интерактивных методов обучения)**

**Лабораторная работа № 1. Реализация матричного умножения (8 часов).** Реализовать алгоритм матричного умножения с помощью парадигмы MapReduce. Составить отчет.

Занятие проводится с использованием метода интерактивного обучения «тематическая дискуссия».

**Лабораторная работа № 2. Решение уравнения теплопроводности с использованием методологии Big Data (8 часов).** Реализовать алгоритм решения уравнения теплопроводности с использованием методологии Big Data. Построить графики решения и ошибки. Составить отчет.

Занятие проводится с использованием метода интерактивного обучения «тематическая дискуссия».

**Лабораторная работа № 3. Реализация алгоритма Брэдли, Файяда и Рейны. (8 часов).** Реализовать алгоритм Брэдли, Файяда и Рейны и выполнить кластеризацию. Анализ кластеризации представить в виде таблиц кластеров, выделенных цветом. Составить отчет.

Занятие проводится с использованием метода интерактивного обучения «тематическая дискуссия».

**Лабораторная работа № 4. Реализация алгоритмов CURE (6 часов).** Реализовать алгоритм CURE и выполнить кластеризацию. Анализ кластеризации представить в виде таблиц кластеров, выделенных цветом. Составить отчет.

Занятие проводится с использованием метода интерактивного обучения «тематическая дискуссия».

**Лабораторная работа № 5. Реализация алгоритма GRGPF (6 часов).** Реализовать алгоритм GRGPF и выполнить кластеризацию. Анализ кластеризации представить в виде таблиц кластеров, выделенных цветом. Составить отчет.

Занятие проводится с использованием метода интерактивного обучения «тематическая дискуссия».

## **V. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ**

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине включает в себя:

- план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;
- требования к представлению и оформлению результатов самостоятельной работы;
- критерии оценки выполнения самостоятельной работы.

### **План-график выполнения самостоятельной работы по дисциплине**

<b>№ п/п</b>	<b>Дата/сроки выполнения</b>	<b>Вид самостоятельной работы</b>	<b>Примерные нормы времени на выполнение</b>	<b>Форма контроля</b>
1	Неделя 1-3	Работа над конспектом лекции, подготовка к лабораторной работе №1	10 часов	ПР-6 (лабораторная работа №1)
2	Неделя 4-7	Работа над конспектом лекции, подготовка к лабораторной работе №2	10 часов	ПР-6 (лабораторная работа №2)
3	Неделя 8-10	Работа над конспектом лекции, подготовка к лабораторной работе №3	10 часов	ПР-6 (лабораторная работа №3)
4	Неделя 11-13	Работа над конспектом лекции, подготовка к лабораторной работе №4	10 часов	ПР-6 (лабораторная работа №4)
5	Неделя 14-17	Работа над конспектом лекции, подготовка к	10 часов	ПР-6 (лабораторная работа №5)

		лабораторной работе №5		
6	Неделя 18	Подготовка к защите лабораторных работ	8 часов	ПР-6 (лабораторная работа №6)
Итого:			58 часа	

### **Характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению**

Самостоятельная работа студентов состоит из подготовки к лабораторным работам в компьютерном классе, работы над рекомендованной литературой и текстами лекций в процессе изучения теоретического материала.

Темы заданий для самостоятельной работы представлены в плане-графике выполнения самостоятельной работы по дисциплине.

При подготовке к лабораторным занятиям необходимо сначала прочитать основные понятия и теоремы по теме. При выполнении задания нужно сначала понять, что требуется в задаче, какой теоретический материал нужно использовать, наметить план решения задачи.

### **Требования к представлению и оформлению результатов самостоятельной работы**

Результатом самостоятельной работы являются отчеты по лабораторным работам. В процессе подготовки отчетов к лабораторным работам у студентов развиваются навыки составления письменной документации и систематизации имеющихся знаний. При составлении отчетов рекомендуется придерживаться следующей структуры:

1. Постановка задачи;
2. Математическая постановка задачи;
3. Описание метода решения;
4. Описание алгоритма метода;
5. Спецификация используемых функций и типов данных;
6. Описание тестов, на которых программа проходит проверку;
7. Анализ результатов численного эксперимента.

### **Критерии оценки выполнения самостоятельной работы**

Отчет по лабораторной работе должен полностью удовлетворять условию задачи. В случае некачественно выполненных отчетов (не соответствующих заявленным требованиям) результирующий балл за работу может быть снижен. Студент должен продемонстрировать отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией. Наличие всех отчетов является допуском к зачету.

Оценка результатов самостоятельной работы организуется как единство двух форм: контроль со стороны преподавателя с использованием рейтинга и



самоконтроль с использованием ЭУК BlackBoard, доступного в компьютерной сети ДВФУ, и содержащего электронные тесты по дисциплине.

### Критерии оценивания лабораторной работы

Результатом лабораторной работы является отчет по лабораторной работе. В процессе подготовки отчетов к лабораторным работам у студентов развиваются навыки составления письменной документации и систематизации имеющихся знаний. При составлении отчетов рекомендуется придерживаться следующей структуры:

1. Постановка задачи;
2. Математическая постановка задачи;
3. Описание метода решения;
4. Описание алгоритма метода;
5. Спецификация используемых функций и типов данных;
6. Описание тестов для проверки работоспособности программы;
7. Результаты численного эксперимента.

Отчет по лабораторной работе должен полностью удовлетворять условию задачи. В случае некачественно выполненных отчетов (не соответствующих заявленным требованиям) результирующий балл за работу может быть снижен. Студент должен продемонстрировать отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией.

## VI. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые модули/разделы / темы дисциплины	Код индикатора достижения компетенции	Результаты обучения	Оценочные средства – наименование	
				текущий контроль	промежуточная аттестация
1	Раздел I. Парадигма Map Reduce	ПК-3.1 Демонстрирует знание методов составления и контроля плана выполняемой научно-исследовательской работы, основы бизнес-планирования	Знает основные методы планирования и контроля НИР в области информационных технологий, основы бизнес-планирования	УО-1; ПР-6; ПР-9	Вопросы к экзамену: 1-5 Курсовой проект
			Умеет строить математические алгоритмы, модели и реализовывать их с помощью языков программирования	УО-1; ПР-6; ПР-9	
			Владеет навыками компьютерной обработки вычислительных задач	УО-1; ПР-6; ПР-9	
		ПК-3.2 Использует методы математического моделирования, принятия решений, разбиения задачи	Знает профессиональную терминологию, содержание ключевых понятий и определений, используемых в теории и практике применения	УО-1; ПР-6; ПР-9	

		на подзадачи, оценивает результат работы команды проекта, риски проекта, составляет бизнес-план	информационных технологий в науке и образовании		
			Умеет применять математический язык, методы при построении моделей объектов профессиональной деятельности с использованием инструментальных средств компьютерного моделирования	УО-1; ПР-6; ПР-9	
			Владеет навыками использования прикладного программного обеспечения для решения задач в профессиональной деятельности	УО-1; ПР-6; ПР-9	
		ПК-3.3 Применяет методы математического моделирования, планирования научно-исследовательской деятельности, работы в научно-исследовательском коллективе	Знает информационные ресурсы и базы данных по научно-исследовательской теме	УО-1; ПР-6; ПР-9	Вопросы к экзамену: 10-15, Курсовой проект
			Умеет самостоятельно расширять и углублять знания в области информационных технологий	УО-1; ПР-6; ПР-9	
			Владеет навыками работы с программными продуктами и информационными ресурсами	УО-1; ПР-6; ПР-9	
2	Раздел 2. Кластеризация больших данных	ПК-3.1 Демонстрирует знание методов составления и контроля плана выполняемой научно-исследовательской работы, основы бизнес-планирования	Знает основные методы планирования и контроля НИР в области информационных технологий, основы бизнес-планирования	УО-1; ПР-6; ПР-9	Вопросы к экзамену: 15-19, Курсовой проект
			Умеет строить математические алгоритмы, модели и реализовывать их с помощью языков программирования	УО-1; ПР-6; ПР-9	
			Владеет навыками компьютерной обработки вычислительных задач	УО-1; ПР-6; ПР-9	
		ПК-3.2 Использует методы математического моделирования, принятия решений, разбиения задачи на подзадачи, оценивает результат работы команды проекта, риски проекта, составляет бизнес-план	Знает профессиональную терминологию, содержание ключевых понятий и определений, используемых в теории и практике применения информационных технологий в науке и образовании	УО-1; ПР-6; ПР-9	Вопросы к экзамену: 20-25, Курсовой проект
			Умеет применять математический язык, методы при построении моделей объектов профессиональной деятельности с использованием инструментальных средств компьютерного моделирования	УО-1; ПР-6; ПР-9	
			Владеет навыками использования прикладного программного обеспечения для решения задач в	УО-1; ПР-6; ПР-9	

		профессиональной деятельности		
	ПК-3.3 Применяет методы математического моделирования, планирования научно-исследовательской деятельности, работы в научно-исследовательском коллективе	Знает информационные ресурсы и базы данных по научно-исследовательской теме	УО-1; ПР-6; ПР-9	Вопросы к экзамену: 26-33, Курсовой проект
		Умеет самостоятельно расширять и углублять знания в области информационных технологий	УО-1; ПР-6; ПР-9	
		Владеет навыками работы с программными продуктами и информационными ресурсами	УО-1; ПР-6; ПР-9	

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также качественные критерии оценивания, которые описывают уровень сформированности компетенций, представлены в разделе VIII.

## **VII. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Основная литература**

1. Data Mining for Service [Electronic resource] / Katsutoshi Yada. — Springer Berlin Heidelberg, 2014. — 291 с. — Режим доступа: <http://lib.dvfu.ru:8080/lib/item?id=chamo:857706&theme=FEFU>
2. Principles of Big Data [Electronic resource] / Jules J. Berman. — Morgan Kaufmann, 2013. — 288 с. — Режим доступа: <http://lib.dvfu.ru:8080/lib/item?id=chamo:809472&theme=FEFU>
3. Воронова Л.И. Big Data. Методы и средства анализа [Электронный ресурс] : учебное пособие / Л.И. Воронова, В.И. Воронов. — Электрон. текстовые данные. — М. : Московский технический университет связи и информатики, 2016. — 33 с. — 2227-8397. — Режим доступа: <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-61463&theme=FEFU>

### **Дополнительная литература**

1. Mining of Massive Datasets [Electronic resource] / Jure Leskovec, Anand Rajaraman, Jeff Ullman. — Cambridge University Press, 2017. — 511 с. <http://www.mmids.org>
2. Афифи А., Эйзен С. Статистический анализ: Подход с использованием ЭВМ. — М.: Мир, 2009. — С. 488. <http://lib.dvfu.ru:8080/lib/item?id=chamo:46908&theme=FEFU>
3. Кобзарь А. И. Прикладная математическая статистика. — М.: Физматлит, 2010.

[http://lib.dvfu.ru:8080/lib/item?id=Lan:/usr/vtls/ChamoHome/visualizer/data\\_lan/data\\_lan+%281905%29.xml&theme=FEFU](http://lib.dvfu.ru:8080/lib/item?id=Lan:/usr/vtls/ChamoHome/visualizer/data_lan/data_lan+%281905%29.xml&theme=FEFU)

4. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Кластерный и факторный анализ. Начальный курс: Учеб. — 7-е изд., испр. — М.: Дело, 2008. <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-4248&theme=FEFU>
5. Факторный, дискриминантный и кластерный анализ: Пер.с англ./Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.; Под ред. И. С. Енюкова. — М.: Финансы и статистика, 2009. — 215 с.  
<http://lib.dvfu.ru:8080/lib/item?id=chamo:309910&theme=FEFU>
6. Экономический факторный анализ: Учебник / Блюмин С.Л., Суханов В.Ф., Чеботарёв С.В., Монография. – Липецк: ЛЭГИ, 2004 – 148 с. ISBN:5-900037-44-4.  
<http://gen.lib.rus.ec/book/index.php?md5=C1DD8B0B7EAB99400E0DEEBF44AC3037>
7. Электронный учебник по статистике. Москва, StatSoft. WEB:  
[www.statsoft.ru/home/textbook/default.htm](http://www.statsoft.ru/home/textbook/default.htm).

#### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Mining massive datasets. <http://www.mmds.org>
2. Big Data от А до Я. Часть 1: Принципы работы с большими данными, парадигма MapReduce. <https://habr.com/company/dca/blog/267361/>
3. Big Data от А до Я. Часть 2: Hadoop. <https://habr.com/company/dca/blog/268277/>
4. Big data от А до Я. Часть 3: Приемы и стратегии разработки MapReduce-приложений. <https://habr.com/company/dca/blog/270453/>
5. Big Data от А до Я. Часть 4: Hbase. <https://habr.com/company/dca/blog/280700/>

#### **Перечень информационных технологий и программного обеспечения**

1. Геоинформационные сервисы <https://habr.com/ru/hub/geo/>
2. ГИС браузер (ArcGIS Online, ArcGIS Explorer, ArcGIS for AutoCAD, ArcGIS для смартфонов и планшетов) <http://introgis.ru/services/sale/freeware/>
3. Пакет программного обеспечения Microsoft Office (Word, Outlook, Power Point, Excel, Photoshop)
4. Пакеты программ ГИС (MapServer, Postgres, PostgreSQL, GRASS GIS, и др.) [http://mapexpert.com.ua/index\\_ru.php?id=75&table=news](http://mapexpert.com.ua/index_ru.php?id=75&table=news)
5. Программные продукты для Windows. Профессиональная ГИС «Панорама» <https://gisinfo.ru/download/download.htm>

## **VIII. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

**Планирование и организация времени, отведенного на изучение дисциплины.** Приступить к освоению дисциплины следует незамедлительно в самом начале учебного семестра. Рекомендуется изучить структуру и основные положения Рабочей программы дисциплины. Обратит внимание, что кроме аудиторной работы (лекции, лабораторные занятия) планируется самостоятельная работа, итоги которой влияют на окончательную оценку по итогам освоения учебной дисциплины. Все задания (аудиторные и самостоятельные) необходимо выполнять и предоставлять на оценку в соответствии с графиком.

В процессе изучения материалов учебного курса предлагаются следующие формы работ: чтение лекций, лабораторные занятия, задания для самостоятельной работы.

*Лекционные занятия* ориентированы на освещение вводных тем в каждый раздел курса и призваны ориентировать студентов в предлагаемом материале, заложить научные и методологические основы для дальнейшей самостоятельной работы студентов.

*Лабораторные занятия* акцентированы на наиболее принципиальных и проблемных вопросах курса и призваны стимулировать выработку практических умений.

Особо значимой для профессиональной подготовки студентов является *самостоятельная работа* по курсу. В ходе этой работы студенты отбирают необходимый материал по изучаемому вопросу и анализируют его. Студентам необходимо ознакомиться с основными источниками, без которых невозможно полноценное понимание проблематики курса.

Освоение курса способствует развитию навыков обоснованных и самостоятельных оценок фактов и концепций. Поэтому во всех формах контроля знаний, особенно при сдаче зачета, внимание обращается на понимание проблематики курса, на умение практически применять знания и делать выводы.

**Работа с литературой.** Рекомендуется использовать различные возможности работы с литературой: фонды научной библиотеки ДВФУ и электронные библиотеки (<http://www.dvfu.ru/library/>), а также доступные для использования другие научно-библиотечные системы.

**Подготовка к экзамену.** К сдаче экзамена допускаются обучающиеся, выполнившие все задания (лабораторные, самостоятельные), предусмотренные учебной программой дисциплины, посетившие не менее 85% аудиторных

занятий.

## IX. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Перечень материально-технического и программного обеспечения дисциплины приведен в таблице.

### Материально-техническое и программное обеспечение дисциплины

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
690001, Приморский край, г. Владивосток, о. Русский, п. Аякс, 10, Корпус 20, ауд. D733	Моноблок lenovo C360G-i34164G500UDK - 13 шт. Мультимедийное оборудование: Экран проекционный ScreenLine Trim White Ice 50 см черная кайма сверху, размер рабочей области 236x147 см Документ-камера Avergence CP355AF ЖК-панель 47", Full HD, LG M4716 CCBA Мультимедийный проектор Mitsubishi EW330U, 3000 ANSI Lumen, 1280x800 Сетевая видеокамера Multipix MP-HD718.	1) Acrobat Pro DC. Договор ЭА-442-15 от 18.01.16 лот 1. Лицензия 20.01.2019. 2) Premiere Elements. Договор ЭА-442-15 от 18.01.16 лот 1. Лицензия 20.01.2019. 3) In Design CC. Договор ЭА-442-15 от 18.01.16 лот 1. Лицензия 20.01.2019. 4) Photoshop CC. Договор ЭА-442-15 от 18.01.16 лот 1. Лицензия 20.01.2019. 5) Academic Campus 500. Договор ЭА-442-15 от 18.01.16 лот 3. Лицензия бессрочно. 6) Academic Reseach. Договор ЭА-442-15 от 18.01.16 лот 3. Лицензия 14.01.2020. 7) Academic Associate Mech. Договор ЭА-442-15 от 18.01.16 лот 3. Лицензия бессрочно.
690001, Приморский край, г. Владивосток, о. Русский, п. Аякс, 10, Корпус 20, ауд. D733a	Компьютер (твердотельный диск - объемом 128 ГБ; жесткий диск - объем 1000 ГБ; форм-фактор - Tower; комплектуется клавиатурой, мышью, монитором AOC i2757Fm; комплектом шнуров эл. питания) модель - M93p1 - 13 шт.	8) SPSS Statistics Premium Campus Edition. Договор ЭА-442-15 от 18.01.16 лот 5. Лицензия бессрочно. 9) SPSS Statistics Premium Base. Договор ЭА-442-15 от 18.01.16 лот 5. Лицензия бессрочно. 10) SPSS Amos. Договор ЭА-442-15 от 18.01.16 лот 5. Лицензия бессрочно.
690001, Приморский край, г. Владивосток, о. Русский, п. Аякс, 10, Корпус 20, ауд. D734	Моноблок HPP-B0G08ES#ACB/8200E AIO i52400S 500G 4.0G 28 PC - 15 шт Мультимедийное оборудование: Экран проекционный ScreenLine Trim White Ice 50 см черная кайма сверху, размер рабочей области 236x147 см Документ-камера Avergence CP355AF ЖК-панель 47", Full HD, LG M4716 CCBA Мультимедийный проектор Mitsubishi EW330U, 3000 ANSI Lumen, 1280x800 Сетевая видеокамера Multipix MP-HD718.	11) АСКОН Компас 3D v17. Поставщик Навиком. Договор 15-03-53 от 20.12.2015. Срок действия договора 31.12.2015. Лицензия бессрочно. 12) Statistica Ultimate Academic Bundle. Договор ЭА-442-15 от 18.01.16 лот 9. Лицензия 14.01.2020. 13) Statistica. Договор ЭА-442-15 от 18.01.16 лот 9. Лицензия 14.01.2020. 14) MathCad Education University Edition. Договор 15-03-49 от 02.12.2015. Лицензия бессрочно. 15) Prompt Translation Server 10 Standart. Договор 15-03-49 от 02.12.2015. Лицензия бессрочно. 16) Prompt Все словари. Договор 15-03-49 от 02.12.2015. Лицензия бессрочно. 17) FineReader 12 Professional Full Academic. Договор 15-03-49 от

		<p>02.12.2015. Лицензия бессрочно.  18) Lingvo x6 Academic Concurrent. Договор 15-03-49 от 02.12.2015. Лицензия бессрочно.  19) Office Professional Plus 2019. Договор № ЭА-261-18. Лицензия .  20) Advanced Threat Analytics Client Management License 2020. Договор № ЭА-261-18. Лицензия 30.06.2020.  21) SQL Server Standard Core 2017. Договор № ЭА-261-18. Лицензия 30.06.2020.  22) Windows Server CAL 2019. Договор № ЭА-261-18. Лицензия 30.06.2020.  23) Windows Server Datacenter Core 2019. Договор № ЭА-261-18. Лицензия 30.06.2020.  24) Windows 10 Enterprise LTSC 2019. Договор № ЭА-261-18. Лицензия 30.06.2020.  25) Windows Edu Per Device 10 Education. Договор № ЭА-261-18. Лицензия 30.06.2020.  26) Autocad 2015. Договор 110002048940. Лицензия 10.09.2020.  27) 3DS MAX 2015. Договор 110002048940. Лицензия 10.09.2020.  28) Alias AutoStudio 2015. Договор 110002048940. Лицензия 12.09.2020.  29) Maya Mental Ray 1 Package 2015. Договор 110002048940. Лицензия 12.09.2020.  30) Turtle For Maya Premium 2015. Договор 110002048940. Лицензия 12.09.2020.  31) MAYA 2015. Договор 110002048940. Лицензия 12.09.2020.  32) MAYA 2016. Договор 110002048940. Лицензия 29.10.2021.  33) Maya Mental Ray 1 Package 2016. Договор 110002048940. Лицензия 29.10.2021.  34) Turtle For Maya Premium 2016. Договор 110002048940. Лицензия 29.10.2021.  35) Plant 3D 2017. Договор 110002048940. Лицензия 30.11.2019.  36) Civil 3D 2017. Договор 110002048940. Лицензия 13.02.2020.  37) Inventor Professional 2017. Договор 110002048940. Лицензия 13.02.2020.  38) Mudbox 2017. Договор 110002048940. Лицензия 13.02.2020.  39) Autocad 2017. Договор 110002048940. Лицензия 16.02.2020.  40) Revit 2017. Договор 110002048940. Лицензия 29.01.2021.  41) 3DS MAX 2020. Договор 110002048940. Лицензия 27.10.2021.  42) AutoCAD 2020. Договор 110002048940. Лицензия 27.10.2021.  43) REVIT 2020. Договор 110002048940.</p>
--	--	--

		Лицензия 27.10.2021. 44) Alias AutoStudio 2020. Договор 110002048940. Лицензия 08.08.2020. 45) MAYA 2020. Договор 110002048940. Лицензия 28.10.2021. 46) Mudbox 2020. Договор 110002048940. Лицензия 29.10.2021. 47) REVIT 2019. Договор 110002048940. Лицензия 28.01.2022. 48) Corel Academic Site. Поставщик Софт Лайн Трейд. Договор ЭА-442-15 от 18.01.16 лот 4. Срок действия договора 30.06.2016. Лицензия закончилась 28.01.2019.
--	--	---

Для проведения учебных занятий по дисциплине, а также для организации самостоятельной работы студентам доступно следующее лабораторное оборудование и специализированные кабинеты, соответствующие действующим санитарным и противопожарным нормам, а также требованиям техники безопасности при проведении учебных и научно-производственных работ.

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья в ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной поддержки.

### **VIII. ФОНДЫ ОЦЕНОЧНЫХ СРЕДСТВ**

Для дисциплины «Компьютерные методы анализа больших данных» используются следующие оценочные средства:

Устный опрос:

1. Собеседование (УО-1)

Письменные работы:

1. Лабораторная работа (ПР-6)

2. Курсовой проект (ПР-9)

**Устный опрос**

Устный опрос позволяет оценить знания и кругозор студента, умение логически построить ответ, владение монологической речью и иные коммуникативные навыки.

Обучающая функция состоит в выявлении деталей, которые по каким-то причинам оказались недостаточно осмысленными в ходе учебных занятий и при подготовке к зачёту.

Собеседование (УО-1) – средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с



изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т.п.

### **Письменные работы**

Письменный ответ приучает к точности, лаконичности, связности изложения мысли. Письменная проверка используется во всех видах контроля и осуществляется как в аудиторной, так и во внеаудиторной работе.

Лабораторная работа (ПР-6) – средство для закрепления и практического освоения материала по определенному разделу.

Курсовой проект (ПР-9) – конечный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве и уровень сформированности аналитических, исследовательских навыков, навыков практического и творческого мышления. Может выполняться в индивидуальном порядке или группой обучающихся

## **Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины**

### **Оценочные средства для промежуточной аттестации**

Промежуточная аттестация студентов по дисциплине «Компьютерные методы анализа больших данных» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной. Форма отчётности по дисциплине – экзамен (1-й, осенний семестр). Экзамен по дисциплине включает ответы на 2 вопроса.

### **Методические указания по сдаче экзамена**

Экзамен принимается ведущим преподавателем. При большом количестве групп у одного преподавателя или при большой численности потока по распоряжению заведующего кафедрой (заместителя директора по учебной и воспитательной работе) допускается привлечение в помощь ведущему преподавателю других преподавателей. В первую очередь привлекаются преподаватели, которые проводили лабораторные занятия по дисциплине в группах.

В исключительных случаях, по согласованию с заместителем директора Школы по учебной и воспитательной работе, заведующий кафедрой имеет право принять экзамен в отсутствие ведущего преподавателя.

Форма проведения экзамена (устная, письменная и др.) утверждается на заседании кафедры по согласованию с руководителем в соответствии с рабочей программой дисциплины.

Во время проведения экзамена студенты могут пользоваться рабочей программой дисциплины, а также с разрешения преподавателя, проводящего экзамен, справочной литературой и другими пособиями (учебниками, учебными пособиями, рекомендованной литературой и т.п.).

Время, предоставляемое студенту на подготовку к ответу на экзамену, должно составлять не более 20 минут. По истечении данного времени студент должен быть готов к ответу.

Присутствие на экзамене посторонних лиц (кроме лиц, осуществляющих проверку) без разрешения соответствующих лиц (ректора либо проректора по учебной и воспитательной работе, директора Школы, руководителя ОПОП или заведующего кафедрой), не допускается. Инвалиды и лица с ограниченными возможностями здоровья, не имеющие возможности самостоятельного передвижения, допускаются экзамен с сопровождающими.

При промежуточной аттестации обучающимся устанавливается оценка «Неудовлетворительно», «Удовлетворительно», «Хорошо» или «Отлично».

В зачетную книжку студента вносится только записи «Удовлетворительно», «Хорошо» или «Отлично», запись «Неудовлетворительно» вносится только в экзаменационную ведомость. При неявке студента на экзамен в ведомости делается запись «не явился».

### **Вопросы к экзамену**

1. Возникновение первых математических понятий.
2. Страны Востока. Египет. Математики Греции. Пифагор. «Начала» Евклида.
3. Творчество Архимеда.
4. Математика Востока.
5. Математика в Европе.
6. Период упадка науки.
7. Эпоха Возрождения.
8. Математика после эпохи Возрождения.
9. Формирование математики переменных величин.
10. Творчество Ньютона и Лейбница.
11. Математика в России.
12. Творчество Ж. Фурье, О. Коши, К. Гаусса, Ан. Пуанкаре.
13. Достижения российской академии наук и российских ученых: П.Л. Чебышева, А.А. Маркова, А.М. Ляпунова.
14. Выдающиеся ученые – А.Н. Тихонов, А.А. Самарский. Математические модели.
15. Модели Солнечной системы. Модели механики сплошной среды.

- Простейшие модели в биологии.
16. Доэлектронная история вычислительной техники. Системы счисления. Абак и счеты.
  17. Первые компьютеры. ENIAC, EDSAC, МЭСМ, М-1. Роль первых ученых - разработчиков компьютеров – Атанасова, Эккерта и Моучли, Дж. фон Неймана, С.А. Лебедева, И.С. Брука.
  18. Поколения ЭВМ. Семейство машин IBM 360/370, машины «Атлас» фирмы ICL, машины фирм Burroughs, CDC, DEC.
  19. Отечественные ЭВМ серий «Стрела», БЭСМ, М-20, «Урал», «Минск». ЭВМ «Сетунь». ЭВМ БЭСМ-6. Семейства ЕС ЭВМ, СМ ЭВМ и «Электроника».
  20. Суперкомпьютеры. ILLIAC IV. Векторно - конвейерные ЭВМ. «Cray-1» и другие ЭВМ Сеймура Крея.
  21. Многопроцессорные ЭВМ классов SMP, MPP, NUMA. Вычислительные кластеры.
  22. Персональные компьютеры и рабочие станции. Микропроцессоры. Роль фирм Apple, IBM, Intel, HP и др.
  23. От сети ARPAnet до Интернета. Локальные вычислительные сети. Сетевые протоколы. Сетевые услуги (удаленный доступ, передача файлов, электронная почта).
  24. Основные области применения компьютеров и вычислительных систем. История математического моделирования и вычислительного эксперимента (Самарский А.А.).
  25. Развитие теории программирования. Библиотеки стандартных программ, ассемблеры (50-е годы XX века).
  26. Языки и системы программирования (60-е годы). Операционные системы (60-70-е годы).
  27. Системы управления базами данных и пакеты прикладных программ (70-80-е годы). Ведущие мировые ученые.
  28. Языки и системы программирования. Первые языки – Фортран, Алгол-60, Кобол. Языки Ada, Pascal, PL/1.
  29. Системы управления базами данных и знаний, пакеты прикладных программ.
  30. Модели данных СУБД. Реляционные и объектно-ориентированные СУБД.
  31. Системы, основанные на знаниях (искусственный интеллект).
  32. Графические пакеты. Машинный перевод.
  33. Программная инженерия. Защита информации.

## Критерии выставления оценки студенту на экзамене

К экзамену допускаются обучающиеся, выполнившие программу обучения по дисциплине, прошедшие все этапы текущей аттестации.

Оценка	Неудовлетворительно	Удовлетворительно	Хорошо	Отлично
Набранная сумма баллов (% выполненных заданий) (макс – 5)	Менее 3 (Менее 60%)	3-3,5 (61-74%)	3,6 -4,4 (75-84%)	4,5-5 (85-100%)

## Оценочные средства для текущей аттестации

Текущая аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация проводится в форме контрольных мероприятий (собеседования, лабораторных работ, курсового проекта) по оцениванию фактических результатов обучения студентов и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- учебная дисциплина (активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);
- степень усвоения теоретических знаний;
- уровень овладения практическими умениями и навыками по всем видам учебной работы;
- результаты самостоятельной работы.

Составляется календарный план контрольных мероприятий по дисциплине. Оценка посещаемости, активности обучающихся на занятиях, своевременность выполнения различных видов заданий ведётся на основе журнала, который ведёт преподаватель в течение учебного семестра.

## Тематика курсовых проектов

1. Разработка хранилища данных результатов переписи
2. Применение методов машинного обучения к большим данным
3. Применение методов Data Mining к большим данным
4. Применение технологии KDD к большим данным
5. Разработка архитектуры и прототипа информационной системы обработки больших данных
6. Применение метода классификации к большим данным

7. Применение нейронных сетей к анализу больших данных
8. Поиск ассоциативных правил в больших данных
9. Исследование и визуализация структуры Facebook, Вконтакте (других социальных сетей)
10. Social media mining
11. Программные средства анализа «больших данных» с открытым исходным кодом.

### Критерии оценки курсовых проектов

Оценка	Требования
<i>«зачтено»</i>	Студент выполняет курсовой проект в полном объёме с соблюдением необходимой последовательности проведения измерений, правильно самостоятельно определяет цель работы; самостоятельно, рационально выбирает необходимое оборудование для получения наиболее точных результатов проводимой работы. Грамотно и логично описывает ход работы, правильно формулирует выводы, точно и аккуратно выполняет все записи, таблицы, рисунки, чертежи, графики, вычисления и т.п., умеет обобщать фактический материал. Допускается два/три недочёта или одна негрубая ошибка и один недочёт. Работа соответствует требованиям и выполнена в срок.
<i>«не зачтено»</i>	Студент выполнил работу не полностью, объём выполненной части не позволяет сделать правильные выводы; не определяет самостоятельно цель работы; в ходе работы допускает одну и более грубые ошибки, которые не может исправить, или неверно производит наблюдения, измерения, вычисления и т.п.; не умеет обобщать фактический материал. Курсовой проект не выполнен.