




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  
  
А.С. Клеин

«УТВЕРЖДАЮ»  
И.о. директора департамента  
  
Еременко А.С.  
« 27 » сентября 2021 г.

### РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Технологии сбора и анализа больших данных

**Направление подготовки - 09.04.01 Информатика и вычислительная техника**  
(Искусственный интеллект и большие данные (совместно с ПАО Сбербанк))

#### Форма подготовки очная

курс 2 семестр 3  
лекции 18 час.  
практические занятия 0 час.  
лабораторные работы 36 час.  
в том числе с использованием МАО практические занятия 0 час  
всего часов аудиторной нагрузки 54 час.  
самостоятельная работа 54 час.  
в том числе на подготовку к экзамену 36 час.  
контрольные работы (количество) не предусмотрены  
курсовая работа / курсовой проект не предусмотрены  
зачет с оценкой не предусмотрен  
экзамен 3 семестр

Рабочая программа дисциплины разработана в соответствии с федеральным государственным образовательным стандартом (ФГОС) высшего образования (ВО) – магистратура по направлению подготовки 09.04.01 Информатика и вычислительная техника, утвержденного приказом Министерства образования и науки Российской Федерации от 19.09.2017 г. № 918 (с изменениями и дополнениями).

Рабочая программа обсуждена на заседании Академии цифровой трансформации протокол № 9.2 от 27 сентября 2022 г.

И.о. директора Академии цифровой трансформации

Еременко А.С.

Составитель (ли): к.т.н. Еременко А.С., Загумённых А.А.

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 курса** – дать слушателям представление о современном положении дел в области сбора и анализа больших данных на основе открытых платформ, а также некоторых облачных решений с использованием этих платформ. Курс направлен как на расширение кругозора в этой области, так и на концептуальное понимание важных теоретических, прикладных и инфраструктурных особенностей современной работы с большими данными с использованием открытых платформ.

### **Задачи курса:**

- научиться производить расчеты с применением технологий анализа больших данных и решать широкий спектр прикладных задач обработки больших наборов данных;
- овладеть навыками количественного и качественного анализа информации при принятии управленческих решений;
- уметь адаптироваться в формируемых организационно-управленческих моделях к конкретным задачам управления.

В результате изучения данной дисциплины у обучающихся формируются следующие профессиональные компетенции:

Код и наименование профессиональной компетенции (результат освоения)	Код и наименование индикатора достижения компетенции
ПК-1 Способен проектировать и разрабатывать системные и прикладные решения по анализу больших данных	ПК-1.1 Владеет инструментарием получения, хранения, передачи и обработки больших данных ПК-1.2 Формулирует и решает системные и прикладные задачи анализа больших данных для конкретных предметных областей ПК-1.3 Способен управлять разработкой продуктов, услуг и решений на основе больших данных
ПК-4 Способен ставить цели и принимать управленческие решения, основанные на анализе больших данных	ПК-4.1 Владеет навыками стратегического управления развитием методологической и технологической инфраструктуры анализа больших данных в организации ПК-4.2 Определяет необходимые ресурсы и инструменты для решения задач с использованием анализа данных; руководит работой команды, вырабатывая командную стратегию на основе анализа данных

Код и наименование индикатора достижения компетенции	Наименование показателя оценивания (результата обучения)
ПК-1.1 Владеет инструментарием получения, хранения, передачи и обработки больших данных	Знает технологии, методы и инструментальные средства обработки больших данных
	Умеет использовать архитектуры и модели баз и хранилищ данных, адаптированные к технологиям больших данных
	Владеет навыками разработки предложений по развитию и совершенствованию системы получения, хранения, передачи, обработки больших данных
ПК-1.2 Формулирует и решает системные и прикладные задачи	Знает существующий опыт разработки и использования продуктов и услуг на основе технологий больших данных

Код и наименование индикатора достижения компетенции	Наименование показателя оценивания (результата обучения)
анализа больших данных для конкретных предметных областей	Умеет разрабатывать проектную и рабочую документацию на разработку аналитических услуг на основе технологий больших данных
	Владеет навыками решения прикладных задач анализа больших данных для конкретных предметных областей
ПК-1.3 Способен управлять разработкой продуктов, услуг и решений на основе больших данных	Знает существующий опыт разработки и использования продуктов и услуг на основе технологий больших данных
	Умеет управлять исполнением проектных работ в области больших данных
	Владеет навыками создания прототипа сервиса на основе аналитики больших данных
	Владеет навыками управления проектными рисками в IT-проекте
ПК-4.1 Владеет навыками стратегического управления развитием методологической и технологической инфраструктуры анализа больших данных в организации	Знает существующие и перспективные методы и программный инструментарий технологий больших данных
	Умеет управлять развитием технологической инфраструктуры анализа больших данных
	Владеет навыками стратегического управления
ПК-4.2 Определяет необходимые ресурсы и инструменты для решения задач с использованием анализа данных; руководит работой команды, выработывая командную стратегию на основе анализа данных	Знает методы создания программного обеспечения для анализа и обработки данных
	Умеет использовать методы проектирования систем анализа и обработки данных
	Владеет навыками работы в распределенных командах

## II. ТРУДОЁМКОСТЬ ДИСЦИПЛИНЫ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ ПО ДИСЦИПЛИНЕ

Общая трудоемкость дисциплины составляет 4 зачётных единиц (144 академических часа).

Видами учебных занятий и работы обучающегося по дисциплине могут являться:

Обозначение	Виды учебных занятий и работы обучающегося
Лаб	Лабораторные работы
Пр	Практические занятия
ОК	Онлайн курс
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	Семестр	Количество часов по видам учебных занятий и работы обучающегося						Формы промежуточной аттестации
			Лек	Лаб	Пр	ОК	СР	Контроль	
1	Введение в проблематику построения анализа больших данных	3	2		0		8		УО-1, УО-4, ПР-11
2	Методы и технологии анализа больших данных	3	8		20		28		
3	Обучение на больших данных	3	8		16		18		
Итого:			18		36		54		Экзамен

### III. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

#### Лекции (18 часов)

#### Тема 1. Введение в проблематику построения систем анализа больших данных (2 часа)

Понятие и классификация больших данных. Размерность и объем данных. «Проклятие размерности». Сценарии применения технологий больших данных в различных отраслях. Жизненный цикл аналитики данных, методологии CRISP-DM и SEMMA. Компетенции и состав команды для работы с данными. Правовые аспекты организации защиты персональных данных. Общий регламент защиты персональных данных (GDPR).

#### Тема 2. Методы и технологии анализа больших данных (8 часов)

Задачи и методы извлечения знаний из больших данных (Data Mining). Инструменты Data Mining. Методы сбора и предобработки данных из различных источников. Примеры применения Data Mining для различных отраслей. Введение в анализ социальных сетей. Базовые алгоритмы на графах и основные возможности графового анализа. Инструменты и методы визуализации графов. Средства визуализации для аналитики данных. Инструментарий для работы с большими данными. Основы работы и функционал компонентов экосистемы Hadoop, парадигма MapReduce. Apache Spark и его компоненты. Обзор облачных платформ для работы с большими данными. Масштабирование и многоуровневое хранение данных: Парадигма NoSQL.

### **Тема 3. Обучение на больших данных (8 часов)**

Особенности представления и обработки символьной и численной информации в нейронных сетях. Методы и средства аннотирования данных для обучения нейронных сетей. Современные архитектуры нейронных сетей: сверточные нейронные сети и автокодировщики, разновидности рекуррентных нейронных сетей, состязательные сети, трансформеры, графовые нейронные сети. Современные алгоритмы и принципы обучения нейронных сетей, особенности обучения сетей различной структуры. Подходы к решению типовых задач, в том числе в области программной инженерии с использованием ИИ. Современные инструментальные средства для разработки и обучения моделей искусственных нейронных сетей. Анализ точности построения и валидация (оценка практической применимости) нейросетевой модели. Системы нечеткого вывода: нечеткие высказывания, вывод в нечеткой логике, правила нечетких продукций. Нечеткие нейронные (гибридные) сети: определение, виды архитектур, обучение и оптимизация.

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

### **Практические занятия (36 час.)**

**Практическая работа 1.** Методы сбора открытых данных из сети интернет (6 часов)

**Практическая работа 2.** Разведочный анализ и визуализация данных (6 часов)

**Практическая работа 3.** Разработка предсказательной модели (6 часов)

**Практическая работа 4.** Анализ поведения пользователей в сети интернет(6 часов)

**Практическая работа 5.** Проектирование и обучение глубокой нейронной сети по распознаванию образов (6 часов)

**Практическая работа 6.** Разработка нечетко-логического регулятора многомерного процесса (6 часов)

## **III. 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</b>	<b>Дата/сроки выполнения</b>	<b>Вид самостоятельной работы</b>	<b>Примерные нормы времени на выполнение</b>	<b>Форма контроля</b>
1	2-3 неделя обучения	Введение в проблематику построения систем. Изучение теоретического материала	8 часов	Собеседование
2	4-12 неделя обучения	Методы и технологии анализа больших данных. Изучение теоретического материала, подготовка практических работ	28 часов	Проверка отчетов, собеседование
3	13-18 неделя обучения	Обучение на больших данных. Изучение теоретического материала, подготовка практических работ.	18 часов	Проверка отчетов, собеседование
Итого:			<b>54 часа</b>	

### **Рекомендации по самостоятельной работе студентов**

Самостоятельная работа обучающихся подразумевает обязательную подготовку к лабораторным занятиям (оформление отчетов), изучение основной и дополнительно литературы по дисциплине, подготовку к текущему контролю и промежуточной аттестации в конце семестра, консультации преподавателей

### **Рекомендации по работе с литературой**

Для более эффективного освоения и усвоения материала рекомендуется ознакомиться с теоретическим материалом по той или иной теме до проведения лабораторного занятия. Всю учебную литературу желательно изучать «под конспект».

Цель написания конспекта по дисциплине – сформировать навыки по поиску, отбору, анализу и формулированию учебного материала.

Работу с теоретическим материалом по теме можно проводить по следующей схеме:

- название темы;
- цели и задачи изучения темы;
- основные вопросы темы;
- характеристика основных понятий и определений, необходимых для усвоения данной темы;

- краткие выводы, ориентирующие на определенную совокупность сведений, основных идей, ключевых положений, систему доказательств, которые необходимо усвоить.

При работе над конспектом обязательно выявляются и отмечаются трудные для самостоятельного изучения вопросы, с которыми уместно обратиться к преподавателю при посещении консультаций, либо в индивидуальном порядке.

### **Подготовка к практическим работам**

Подготовку к каждой практической работе или к практическому занятию каждый студент должен начать с изучения теоретического материала и ознакомления с планом, который отражает содержание предложенной темы. Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса. Результат такой работы должен проявиться в способности студента свободно ответить на теоретические вопросы по теме задания, правильном его выполнении.

В процессе выполнения практической работы студент должен создать требуемый документ с помощью предлагаемого программного средства и выполнить требуемые в задании операции. Задание по лабораторной или практической работе содержит методические указания по подготовке документа, который должен быть получен в результате выполнения работы. При подготовке следует их внимательно прочесть.

Отчет по практической работе оформляется и отчитывается в электронном виде: формат листа А4, книжная ориентация страницы. Отчеты по всем лабораторным работам имеют единый титульный лист, на котором указывается наименование дисциплины, ФИО и группа исполнителя, ФИО преподавателя, принимающего отчеты. В отчете по каждой лабораторной работе должно быть представлено наименование работы, цель, ход выполнения работы (скриншоты, краткое текстовое описание), выводы по результатам работы.

### **Критерии оценки практических работ**

– 100-86 - выполнены все задания практической работы, студент четко и без ошибок ответил на все контрольные вопросы.

– 85-76 - выполнены все задания практической работы; студент ответил на все контрольные вопросы с замечаниями.

– 75-61 выполнены все задания практической работы с замечаниями; студент ответил на все контрольные вопросы с замечаниями.



- 60-50 баллов - студент не выполнил или выполнил неправильно задания практической работы; студент ответил на контрольные вопросы с ошибками или не ответил на контрольные вопросы.

#### IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы/темы дисциплины	Код индикатора достижения компетенции	Результаты обучения	Оценочные средства - наименование	
				текущий контроль	промежуточная аттестация
1.	Тема 1 Введение в проблематику построения систем анализа больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1,	Экзамен Вопросы 1-5
			Умеет		
			Владеет		
2.	Тема 2. Методы и технологии анализа больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1, Практические работы 1-4	Экзамен Вопросы 6-15
			Умеет		
			Владеет		
3.	Тема 3. Обучение на больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1, Практические работы 5-6	Экзамен Вопросы 16-25
			Умеет		
			Владеет		

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.

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

### **Основная литература**

*(электронные и печатные издания)*

1. Виктор Майер-Шенбергер Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / Виктор Майер-Шенбергер, Кеннет Кукьер. — Москва : Манн, Иванов и Фербер, 2014. — 220 с. — ISBN 978-5-91657-936-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/39150.html>

2. Воронов В.И. Data Mining - технологии обработки больших данных : учебное пособие / Воронов В.И., Воронова Л.И., Усачев В.А.. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/81324.html>

3. Железнов М.М. Методы и технологии обработки больших данных : учебно-методическое пособие / Железнов М.М.. — Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/101802.html>

4. Келлехер Д. Наука о данных: базовый курс: учебное пособие/ науч. ред. З. Мамедьяров; пер. с англ. М. Белоголовского. — Москва: Аль пина Паблишер, 2020. — 224 с. - URL:

### **Дополнительная литература**

*(печатные и электронные издания)*

1. Болотова Л.С. Системы искусственного интеллекта: модели и технологии, основанные на знаниях: учебник / ФГБОУ ВПО РГУИТП; ФГАУ ГНИИ ИТТ «Информика». — М.: Финансы и статистика, 2012. — 664 с.: ил. <http://www.studentlibrary.ru/book/ISBN9785279035304.html>

2. Загоруйко Н.Г. Когнитивный анализ данных. — Новосибирск: Гео, 2013. — 183 с. ISBN 978-5-906284-04-4.

3. Осипов Г.С. Методы искусственного интеллекта. — М.: ФИЗМАТЛИТ, 2011. — 296 с. <http://www.studentlibrary.ru/book/ISBN9785922113236.html>

4. Низаметдинов Ш.У., Румянцев В.П. Анализ данных: учебное пособие. — М.: НИЯУ «МИФИ», 2012. — 288 с. ISBN 978-5-7262-1687- <http://znanium.com/bookread2.php?book=567083>

5. Кухаренко Б.Г. Интеллектуальные системы и технологии – М.: Московская государственная академия водного транспорта, 2015. – 116 с. и <http://www.iprbookshop.ru/47933.html>

6. Павлов С.Н. Системы искусственного интеллекта. Часть 1 [Электронный ресурс]: учебное пособие. – Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2011. – 176 с. <http://www.iprbookshop.ru/13974.html>

7. Павлов С.Н. Системы искусственного интеллекта. Часть 2 [Электронный ресурс]: учебное пособие. – Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2011. – 194 с. <http://www.iprbookshop.ru/13975.html>

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

1. <http://machinelearning.ru/> MachineLearning.ru Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных.
2. <http://www.inftech.webservis.ru/it/database/datamining/ar2.html> Дюк В.А. Data Mining – интеллектуальный анализ данных.
3. <http://kek.ksu.ru/EOS/dm.pdf> Степанов Р.Г. Технология Data Mining: Интеллектуальный анализ данных / Казань, 2008.

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

Для проведения практических занятий требуется следующее свободно распространяемое программное обеспечение: Microsoft Excel, Google Colaboratory, Python.GNU R, Python, а также автоматическая тестирующая система CATS ДВФУ <https://imcs.dvfu.ru/cats/>.

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

**Планирование и организация времени, отведенного на изучение дисциплины.** Приступить к освоению дисциплины следует незамедлительно в самом начале учебного семестра. Рекомендуются 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

итогах 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.

*Практические 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

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 (практические работы, самостоятельные), предусмотренные учебной программой дисциплины, посетившие не менее 85% аудиторных занятий.

## VII. 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
<p>690922, Приморский край, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, корпус D, ауд. D 733,733а. Учебная аудитория для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации</p>	<p>Помещение укомплектовано специализированной учебной мебелью (посадочных мест – 13) Оборудование: ЖК-панель 47", Full HD, LG M4716 CCBA – 1 шт. Доска аудиторная, Моноблок Lenovo C360G-i34164G500UDK с лицензионными программами Microsoft Office 2013(13 шт.) и аудиовизуальными средствами проектор Panasonic DLPProjectorPT-D2110XE</p>	<p>1С Предприятия8 (8.2), 7-Zip, ABBYY Lingvo12,Alice 3, Anaconda3,Autodesk,CodeBlocks,CorelDRAW X7,Dia,Directum4.8,DosBox-0.74,Farmanager,Firebird 2.5,FlameRobin,Foxit Reader,Free Pascal,Geany,Ghostscript,Git,Greenfoot,gsview,Inscapе0.91,Java,Java development Kit,Kaspersky,Lazarus,LibreOffice4.4,MatLab R2017b,Maxima 5.37.2,Microsoft Expression,Microsoft Office 2013,Microsoft Silverlight,Microsoft Silverlight 5SDK-русский,MicrosoftSistem Center,Microsoft Visual Studio 2012,MikTeX2.9,MySQL,NetBeans,Notepad++,Oracle VM VirtualBox,PascalABC.NET,PostgreSQL 9.4,PTC Mathcad,Putty,PyQt GPL v5.4.1 for Pythonv 3.4,Python2.7(3.4,3.6),QGIS Brighton,RStudio,SAM CoDeC Pack,SharePoint,Strawberry Perl,Tecnomatix,TeXnicCenter,TortoiseSVN,Unity2017.3.1f1,Veusz,Vim8.1,Visual Paradigm CE,Visual Studio2013,Windows Kits,Windows Phone SDK8.1,Xilinx Design ToolsAcrobat ReaderDC,AdobeBridge CS3,AdobeDeviceCentralCS3,Adobe ExtendScript Toolkit 2,Adobe Photoshpe CS3,DVD-студия Windows,GoogleChrome,Internet Explorer,ITMOproctor,Mozilla Firefox, Visual Studio Installer,Windows Media Center, WinSCP,</p>

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

№ п/п	Контролируемые разделы/темы дисциплины	Код индикатора достижения компетенции	Результаты обучения	Оценочные средства - наименование	
				текущий контроль	промежуточная аттестация
1.	Тема 1 Введение в проблематику построения систем анализа больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1,	Экзамен Вопросы 1-5
			Умеет		
			Владеет		
2.	Тема 2. Методы и технологии анализа больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1, Практические работы 1-4 (ПР-4) Кейс (ПР-2)	Экзамен Вопросы 6-15
			Умеет		
			Владеет		
3.	Тема 3. Обучение на больших данных	ПК – 1.1 ПК- 1.2 ПК – 1.3 ПК – 4.1 ПК – 4.2	Знает	Собеседование УО1, Практические работы 5-6 (ПР-4) Кейс (ПР-2)	Экзамен Вопросы 16-25
			Умеет		
			Владеет		

Используются следующие оценочные средства:

Устный опрос:

1. Собеседование (УО-1)

Письменные работы:

1. Практическая работа (ПР-4)

2. Кейс (ПР-2)

**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**

Письменная проверка используется во всех видах контроля и осуществляется как в аудиторной, так и во внеаудиторной работе.

Кейс (ПР-2) – проблемное задание, в котором обучающемуся предлагают осмыслить реальную профессионально-ориентированную ситуацию, необходимую для решения данной проблемы..

Практическая работа (ПР-4) – средство для закрепления и практического освоения материала по определенному разделу.

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

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

Промежуточная аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной. Форма отчётности по дисциплине – экзамен (3-й семестр).

Экзамен проводится в письменной форме с использованием защиты проекта.

### **Критерии выставления оценки студенту**

<b>Баллы (рейтинговой оценки)</b>	<b>Оценка зачета/ экзамена (стандартная)</b>	<b>Требования к сформированным компетенциям</b>
86-100	«зачтено»/ «отлично»	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.
76-85	«зачтено»/ «хорошо»	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.

Баллы (рейтинговой оценки)	Оценка зачета/ экзамена (стандартная)	Требования к сформированным компетенциям
61-75	«зачтено»/ «удовлетворительно»	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.
0-60	«не зачтено»/ «неудовлетворительно»	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

### Вопросы к экзамену

1. Понятие и классификация больших данных. Размерность и объем данных. «Проклятие размерности».
2. Понятие и формирование озера данных Data Lake.
3. Методология исследования данных CRISP-DM. Задачи каждой фазы и возможные ошибки.
4. Методология исследования данных SEMMA. Отличия от CRISP-DM.
5. Компетенции и состав команды для работы с данными.
6. Правовые аспекты организации защиты персональных данных регламент защиты персональных данных (GDPR).
7. Задачи и методы извлечения знаний из больших данных.
8. Методы сбора данных из различных источников.
9. Методы предобработки данных.
10. Анализ социальных сетей.
11. Визуализация данных.
12. Назначение и характеристика компонент экосистемы Hadoop.
13. Парадигма MapReduce.
14. Apache Spark и его компоненты.
15. Облачные платформы для работы с большими данными.
16. Масштабирование и многоуровневое хранение данных: Парадигма NoSQL.
17. Особенности представления и обработки символьной и численной



информации в нейронных сетях.

18. Методы и средства аннотирования данных для обучения нейронных сетей.

19. Современные алгоритмы и принципы обучения нейронных сетей, особенности обучения сетей различной структуры

20. Методы оценки качества нейросетевых моделей.

21. Валидация и интерпретируемость результатов работы нейросетевых моделей.

22. Системы нечеткого вывода: нечеткие высказывания, вывод в нечеткой логике, правила нечетких продукций.

23. Понятие нечеткой нейронной сети: определение и виды архитектур.

24. Обучение и алгоритмы оптимизации нечеткой нейронной сети.

25. Центры обработки данных: назначение, оборудование, уровни

### **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, практических работ, решения тестов) по оцениванию фактических результатов обучения студентов.

#### **Отчет по практической работе**

Содержание отчета: 1. Цель и задачи лабораторной работы. 2. Примеры запросов через API с описанием используемых параметров. 3. Примеры ответов по запросам и анализ полученной информации 4. Описание этапов парсинга сайта с указанием проводимых настроек 5. Анализ полученной информации. 6. Обосновать выбор модуля для ИИ. 7. Описать выявленные зависимости в данных.

Отчет по практической работе представляется в электронном виде. Защита отчета проходит в форме доклада студента по выполненной работе и ответов на вопросы преподавателя. В случае, если оформление отчета и поведение студента во время защиты соответствуют указанным требованиям, студент получает максимальное количество баллов.

Основаниями для снижения количества баллов в диапазоне от max до min являются:

- небрежное выполнение,
- отсутствие выводов,
- нарушение сроков предоставления отчета.

Отчет не может быть принят и подлежит доработке в случае:

- серьезного несоответствия техническому заданию;
- отсутствия минимально необходимого количества тестовых примеров;
- некорректной работы программы и т.п.

### **Кейс/ситуационное задание**

Описание технологии применения кейса: задание выдается и выполняется во время лекционных занятий после изложения соответствующего теоретического материала; форма проведения - групповой анализ конкретной ситуации, время, отводимое на выполнение задания, - 30 минут. Форма представления результатов — устный доклад и письменный отчет по результатам обсуждений.

### **Примеры кейсовых заданий**

*Тема 1. Введение в проблематику построения систем анализа больших данных*

Задания для решения кейса/комплект:

Кейс 1. Подготовка плана проекта по анализу данных

Требуется представить план проекта по анализу данных на заданную тему (например, прогнозирование оттока клиентов фирмы). Работа состоит из следующих этапов: определение целей проекта и требований со стороны заказчика; определение цели анализа данных; разработка плана проекта; распределение обязанностей участникам проекта в зависимости от выполняемой роли.

*Тема 2 Методы и технологии анализа больших данных*

Кейс 2. Разработка требований заказчика к информационно-аналитической системе

Цель: освоение базовых инструментов составления требований заказчика. Требуется описать в 3-5 пользовательских историях и 1-2 сценариях использования любое приложение на ноутбуке или мобильном телефоне. Приветствуется добавление нового функционала.

Кейс 3. Проведение переговоров с подрядчиками и поставщиками оборудования и программного обеспечения.

Требуется определить требования и составить план переговоров с подрядчиками и поставщиками оборудования и программного обеспечения