



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

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

«СОГЛАСОВАНО»

Руководитель ОП

Артемяева И.Л.

« 26» января 2022 г.

«Утверждаю»

И.о. директора департамента

Смагин С.В.

« 26» января 2022 г.



**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Интеллектуальный анализ данных

**Направление подготовки – 09.04.04 «Программная инженерия»**

(Разработка программно-информационных систем)

**Форма подготовки очная**

курс 2 семестр 3  
лекции 18 час.  
практические занятия 0 час.  
лабораторные работы 18 час.  
в том числе с использованием МАО лек. 0 / пр. 0 / лаб. 18 час.  
всего часов аудиторной нагрузки – 36 час.  
в том числе с использованием МАО – 18 час.  
самостоятельная работа 72 час  
в том числе на подготовку к экзамену 0 час.  
контрольные работы (количество) не предусмотрены  
курсовая работа / курсовой проект не предусмотрен  
зачет 3 семестр  
экзамен не предусмотрен

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки 09.04.04 **Программная инженерия**, утвержденного приказом Министерства образования и науки РФ от 19.09.2017 № 932 (с изменениями и дополнениями).

Рабочая программа обсуждена на заседании департамента программной инженерии и искусственного интеллекта, протокол № 1.1 от «24» января 2022 г.

И.о. директора департамента программной инженерии и искусственного интеллекта к.т.н. Смагин С.В.

Составитель (ли): доцент департамента ПИИИИ к.т.н. Смагин С.В.

Владивосток  
2022

**Оборотная сторона титульного листа РПД**

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

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

Директор департамента \_\_\_\_\_

(подпись) (И.О.Фамилия)

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

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

Директор департамента \_\_\_\_\_

(подпись) (И.О.Фамилия)

## 1. Цели и задачи освоения дисциплины:

**Цель:** изучение современных методов решения задач интеллектуального анализа экспериментальных данных, получаемых в различных областях науки, экономики и бизнеса, освоение технологий оценки свойств этих методов и применения их на практике.

### **Задачи:**

1. Изучение основных понятий интеллектуального анализа данных.
2. Изучение классификации моделей предметных областей, критериев их построения, анализа и сравнения.
3. Рассмотрение постановок задач индуктивного формирования баз знаний для проблемно-независимых и проблемно-ориентированных моделей предметных областей, а также рассмотрение основных методов для решения этих задач.
4. Изучение подходов к организации, проведению и интерпретации результатов экспериментов для оценки внешних и внутренних свойств методов интеллектуального анализа данных на модельных и реальных данных.
5. Проведение экспериментов по оценке внешних и внутренних свойств метода направленного поиска на примере упрощенной онтологии медицинской диагностики в программе Microsoft Excel и среде Google Colaboratory.

Для успешного изучения дисциплины у обучающихся должны быть сформированы предварительные компетенции:

- способность к самоорганизации и самообразованию;
- владение основными концепциями, принципами, теориями и фактами, связанными с информатикой;
- владение основными вычислительными алгоритмами решения оптимизационных задач;
- владение вероятностными методами моделирования данных и принятия решений.

Профессиональные компетенции выпускников и индикаторы их достижения:

Тип задач	Код и наименование профессиональной компетенции (результат освоения)	Код и наименование индикатора достижения компетенции
Производственно-технологический	ПК-4. Способен создавать программное обеспечение для анализа и обработки информации.	ПК-4.1. Демонстрирует знание методов создания программного обеспечения для анализа и обработки информации.
		ПК-4.2. Использует методы создания программного обеспечения для анализа и обработки информации.
		ПК-4.3. Применяет методы создания программного обеспечения для анализа и обработки информации, требуемых в профессиональной деятельности.

Код и наименование индикатора достижения компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
ПК-4.1. Демонстрирует знание методов создания программного обеспечения для анализа и обработки информации.	Знает внешние и внутренние свойства алгоритмов обработки больших объемов данных, а также основные достижения современных технологий в этой области.
	Умеет сформулировать основные требования к свойствам алгоритма интеллектуального анализа данных.
	Владеет схемой компьютерного эксперимента для вычисления оценок свойств алгоритмов интеллектуального анализа данных.
ПК-4.2. Использует методы создания программного обеспечения для анализа и обработки информации.	Знает критерии выбора наиболее подходящего метода распознавания и обработки данных для произвольной предметной области.
	Умеет провести компьютерный эксперимент для вычисления оценок свойств алгоритмов интеллектуального анализа данных.
	Владеет способностью сравнить два алгоритма на основе их внешних и внутренних оценок.
ПК-4.3. Применяет методы создания программного обеспечения для анализа и обработки информации, требуемых в профессиональной деятельности.	Знает современные платформы, среды, программные средства и библиотеки, в которых реализованы наиболее популярные и эффективные алгоритмы интеллектуального анализа данных.
	Умеет оценить алгоритм интеллектуального анализа данных с точки зрения требуемого времени на его построение и качества его работы.
	Владеет практическими навыками работы в Microsoft Azure и Google Colaboratory.

## 2.Трудоёмкость дисциплины и видов учебных занятий по дисциплине

Общая трудоёмкость дисциплины составляет 3 зачётных единиц (108 академических часов).

(1 зачетная единица соответствует 36 академическим часам)

Видами учебных занятий и работы обучающегося по дисциплине могут являться:

Обозначение	Виды учебных занятий и работы обучающегося
Лек	Лекции
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	Основные понятия и примеры прикладных задач	3	2	2					УО-1, УО-4, ПР-6 экзамен
2	Изучение критериев построения, анализа и сравнения математических моделей предметных областей	3	2	2					
3	. Способы формирования баз знаний и их сравнение	3	2	2					
4	Свойства методов интеллектуального анализа данных и различия существующих подходов к оцениванию их свойств	3	2	2			72		
5	Организация и проведение экспериментов на модельных и реальных данных	3	4	6					
6	Проведение экспериментов по оценке внешних и внутренних свойств методов интеллектуального анализа данных на примере предметной области медицинской диагностики	3	6	4					
	Итого:		18	18			72		экзамен

### I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

**Трудоемкость теоретической части курса 18 час.**

**Лекции (18 час)**

#### **Тема 1. Основные понятия и примеры прикладных задач (2 час.)**

Рассматриваются понятия: объекты; признаки; свойства; классы; типы шкал; модель предметной области; задача обучения по прецедентам; обучающая и контрольная выборки; модельные и реальные данные; метод решения; алгоритм обучения; база знаний; критерии качества обучения; интерпретируемость базы знаний; примеры прикладных задач.

## **Тема 2. Изучение критериев построения, анализа и сравнения математических моделей предметных областей (2 час.)**

Рассматриваются понятия: математическая модель предметной области; виды математических моделей предметных областей (проблемно-независимые и проблемно-ориентированные); свойства моделей предметных областей; исследование свойств моделей предметных областей; критерии выбора моделей и методы отбора признаков; прямая и обратная задачи математического моделирования; критерии сравнения моделей.

## **Тема 3. Способы формирования баз знаний и их сравнение (2 час.)**

Рассматриваются понятия: экспертный способ; автоматический (индуктивный) способ; достоинства и недостатки каждого способа; алгоритмы индуктивного формирования баз знаний; традиционные и новые постановки задач классификации и диагностики.

## **Тема 4. Свойства методов интеллектуального анализа данных и различия существующих подходов к оцениванию их свойств (2 час.)**

Рассматриваются понятия: результат работы (детерминированный и вероятностный); объяснение результатов (хорошо интерпретируемое и плохо интерпретируемое); характеристики получения результатов; отличия методов интеллектуального анализа данных от методов машинного обучения.

## **Тема 5. Организация и проведение экспериментов на модельных и реальных данных (4 час.)**

Рассматриваются понятия: внешние и внутренние свойства методов интеллектуального анализа данных; методика экспериментального исследования свойств методов интеллектуального анализа данных на модельных и реальных данных; сравнение методов интеллектуального анализа данных; соревнования среди алгоритмов обучения.

## **Тема 6. Проведение экспериментов по оценке внешних и внутренних свойств методов интеллектуального анализа данных на примере предметной области медицинской диагностики (6 час.)**

Рассматриваются понятия: упрощенная модель предметной области медицинской диагностики; модельная база знаний; параметры и неизвестные модели; индуктивно формируемая база знаний; формирование оценок внешних и внутренних свойств в программе Microsoft Excel и среде Google Colaboratory; заключение о качестве метода на основе оценки его свойств.

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

### **Практические занятия (0 час.)**

## **Лабораторные работы (18 час.)**

### **Лабораторная работа 1. Пример прикладной задачи (2 час.)**

Задание: рассмотреть предметную область медицинской диагностики на примере ее упрощенной модели; проанализировать существующие в области объекты; выделить признаки объектов; указать типы шкал выделенных признаков; сформировать контекстные условия их совместного использования.

### **Лабораторная работа 2. Критерии построения и анализа математических моделей предметных областей (2 час.)**

Задание: на основе решения предыдущего задания сформировать постановку задачи медицинской диагностики; проанализировать и оценить свойства модели; предложить метод решения задачи.

### **Лабораторная работа 3. Способы формирования баз знаний (2 час.)**

Задание: на основе решения предыдущего задания выделить в модели термины для описания знаний и термины для описания действительности; сформировать соглашения, связывающие эти термины; сформировать значения базы знаний экспертным способом.

### **Лабораторная работа 4. Разработка методов интеллектуального анализа данных (2 час.)**

Задание: на основе решения предыдущего задания и методов случайного и направленного поиска разработать аналогичные методы для решения задачи медицинской диагностики в рамках ее упрощенной модели.

### **Лабораторная работа 5. Проведение экспериментов на модельных данных (6 час.)**

Задание: на основе решения предыдущего задания получить оценку внешних и внутренних свойств разработанных методов интеллектуального анализа данных в рамках проведения компьютерных экспериментов на модельных данных в программе Microsoft Excel и среде Google Colaboratory.

### **Лабораторная работа 6. Анализ оценок внешних и внутренних свойств методов интеллектуального анализа данных (4 час.)**

Задание: на основе решения предыдущего задания визуализировать полученные оценки внешних и внутренних свойств разработанных методов интеллектуального анализа данных в программе Microsoft Excel и среде Google Colaboratory; сделать заключение о качестве методов на основе сравнения оценок их свойств.

## **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

/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1	2-4 неделя обучения	Рассмотреть предметную область медицинской диагностики на примере ее упрощенной модели	8 часов	Собеседование
2	5-6 неделя обучения	Проанализировать и оценить свойства модели, предложить метод решения задачи	8 часов	Собеседование
3	7-8 неделя обучения	Выделить в модели термины для описания знаний и термины для описания действительности, сформировать соглашения, связывающие эти термины, сформировать значения базы знаний экспертным способом	16 часов	Проверка отчетов, собеседование
4	9-10 неделя обучения	Разработать аналоги методов случайного и направленного поиска для решения задачи медицинской диагностики в рамках ее упрощенной модели	16 часов	Проверка отчетов, собеседование
5	11-17 неделя обучения	Получить оценку внешних и внутренних свойств разработанных методов в рамках проведения компьютерных экспериментов на модельных данных, визуализировать полученные оценки, сделать заключение о	24 часа	Проверка отчетов, собеседование



		качестве методов на основе сравнения оценок их свойств		
Итого:			72 часа	

### **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 лабораторным работам**

Подготовку к каждой лабораторной 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 лабораторной работы или практического задания

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.

### **Критерии оценки лабораторных работ**

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

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

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

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

### **Подготовка презентации и доклада**

Для подготовки презентации рекомендуется использовать: PowerPoint, MS Word, Acrobat Reader, LaTeX-овский пакет beamer. Последовательность подготовки презентации:

1. Четко сформулировать цель презентации: вы хотите свою аудиторию мотивировать, убедить, заразить какой-то идеей или просто формально отчитаться.

2. Определить каков будет формат презентации: живое выступление (тогда, сколько будет его продолжительность) или электронная рассылка (каков будет контекст презентации).

3. Отобрать всю содержательную часть для презентации и выстроить логическую цепочку представления.

4. Определить ключевые моменты в содержании текста и выделить их.

5. Определить виды визуализации (иллюстрации, образы, диаграммы, таблицы) для отображения их на слайдах в соответствии с логикой, целью и спецификой материала.

6. Подобрать дизайн и форматировать слайды (количество картинок и текста, их расположение, цвет и размер).

7. Проверить визуальное восприятие презентации.

***Практические советы по подготовке презентации*** - готовьте отдельно:

- печатный текст + слайды + раздаточный материал;

- *слайды* – визуальная подача информации, которая должна содержать минимум текста, максимум изображений, несущих смысловую нагрузку, выглядеть наглядно и просто;
- *текстовое содержание презентации* – устная речь или чтение, которая должна включать аргументы, факты, доказательства и эмоции;
- *рекомендуемое число слайдов* 17-22;
- *обязательная информация для презентации*: тема, фамилия и инициалы выступающего; план сообщения; краткие выводы из всего сказанного; список использованных источников;
- *раздаточный материал* – должен обеспечивать ту же глубину и охват, что и живое выступление: люди больше доверяют тому, что они могут унести с собой, чем исчезающим изображениям, слова и слайды забываются, а раздаточный материал остается постоянным осязаемым напоминанием; раздаточный материал важно раздавать в конце презентации; раздаточный материалы должны отличаться от слайдов, должны быть более информативными.

	50-60 баллов (неудовлетворительно)	61-75 баллов (удовлетворительно)	76-85 баллов (хорошо)	86-100 баллов (отлично)
критерии	Содержание критериев			
Раскрытие проблемы	Проблема не раскрыта. Отсутствуют выводы	Проблема раскрыта не полностью. Выводы не сделаны и/или выводы не обоснованы	Проблема раскрыта. Проведен анализ проблемы без привлечения дополнительной литературы. Не все выводы сделаны и/или обоснованы	Проблема раскрыта полностью. Проведен анализ проблемы с привлечением дополнительной литературы Выводы обоснованы
Представление	Представляемая информация логически не связана. профессиональные термины Студент демонстрирует неумение использовать понятийный аппарат	Представляемая информация не систематизирована и/или не последовательна. Студент демонстрирует затруднения с использованием научно-понятийного аппарата и терминологии учебной дисциплины;	Представляемая информация не систематизирована и последовательна. Студент демонстрирует умение пользоваться концептуально-понятийным аппаратом в процессе анализа основных проблем в рамках данной темы	Представляемая информация систематизирована, последовательна и логически связана. Студент демонстрирует отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией соответствующей научной области.
Ответы на вопросы	Нет ответов на вопросы	Только ответы на элементарные вопросы	Ответы на вопросы полные и/или частично полные	Ответы на вопросы полные, с приведением примеров и/или пояснений
Оформление	Не использованы технологии Power Point. Больше 4 ошибок в представляемой информации	Использованы технологии Power Point частично. 3-4 ошибки в представляемой информации	Использованы технологии Power Point. Не более 2 ошибок в представляемой информации	Широко использованы технологии (Power Point и др.).Отсутствуют ошибки в представляемой информации

## Критерии оценки презентации доклада

Менее 60 баллов	незачтено	неудовлетворительно
От 61 до 75 баллов	зачтено	удовлетворительно
От 76 до 85 баллов	зачтено	хорошо
От 86 до 100 баллов	зачтено	отлично

### IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы/темы дисциплины	Код индикатора достижения компетенции	Результаты обучения	Оценочные средства	
				текущий контроль	промежуточная аттестация
1.	Темы 1-9	ПК-4.1. Демонстрирует знание методов создания программного обеспечения для анализа и обработки информации.	Знает способы получения оценок моделей и методов машинного обучения, а также критериев их сравнения.	Собеседование УО1, круглый стол УО-4	Экзамен Вопросы 1-13
			Умеет поставить задачу машинного обучения в заданной предметной области и предложить метод ее решения.		
			Владеет навыками работы с платформами, средами и программными средствами, в библиотеках которых реализованы методы машинного обучения.		
		ПК-4.2. Использует методы создания программного обеспечения для анализа и обработки информации.	Знает способы выбора наиболее подходящего метода машинного обучения для заданной предметной области на основе оценки и анализа свойств существующих методов.		

			<p>Умеет проанализировать заданную предметную область на предмет возможности и необходимости применения в ней методов машинного обучения.</p> <p>Владеет навыками применения заданного метода машинного обучения (на основе существующих библиотек), а также получения результата (на основе заданного набора данных).</p>		
		ПК-4.3. Применяет методы создания программного обеспечения для анализа и обработки информации, требуемых в профессиональной деятельности.	Знает основные репозитории, в которых хранятся наборы данных для машинного обучения, и способы их использования.	Л/работы 1-4, ПР-6	
			Умеет обработать заданный набор данных различными методами машинного обучения и визуализировать результат.		
			Владеет способностью интерпретации полученных результатов и их улучшения посредством настройки значений параметров.		
2.	Темы 10-15	ОПК-7.1. Демонстрирует знание методов и средств получения, хранения, переработки и трансляции информации	<p>Знает области применения, преимущества и недостатки основных алгоритмов и технологий обработки данных.</p> <p>Умеет оценить</p>	Собеседование УО1, круглый стол УО-4	Экзамен Вопросы 14-20 (практическая часть)

		<p>посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях для решения задач в области профессиональной деятельности.</p>	<p>степень применимости, а также временную и вычислительную сложности выбранного алгоритма для решения конкретной прикладной задачи.</p>		
		<p>ОПК-7.2. Использует методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях.</p>	<p>Владеет способностью провести анализ степени применимости имеющихся алгоритмов для всех возможных (альтернативных) решений конкретной прикладной задачи.</p>		
			<p>Знает критерии для выбора наиболее подходящего алгоритма и технологии обработки данных для произвольной предметной области.</p>	<p>Л/работы 5-8, ПР-6</p>	
			<p>Умеет оценить временные и ресурсные затраты на решение конкретной прикладной задачи выбранным алгоритмом обработки данных.</p>		
		<p>ОПК-7.3. Применяет методы и средства получения,</p>	<p>Владеет способностью выбрать из имеющегося набора алгоритмов одного, наиболее подходящего с точки зрения экономии ресурсов, а также качества получаемого результата.</p>		
			<p>Знает методы и средства сбора, проверки и систематизации информации для</p>		

		хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях при решении профессиональных задач.	произвольной предметной области. Умеет подготовить единообразный набор данных и провести его обобщение (с записью результата в базу знаний) для произвольной предметной области. Владеет способностью оценить базу знаний с точки зрения оценки ее свойств адекватности, полноты, сложности, непротиворечивости и т.д.		
--	--	---	--	--	--

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. 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 978-5-7262-1687-4.



<http://znanium.com/bookread2.php?book=567083>

**Дополнительная литература**  
(печатные и электронные издания)

1. Вагин В.Н. Достоверный и правдоподобный вывод в интеллектуальных системах: учебное пособие. – М.: Физматлит, 2008. – 704 с.  
<http://znanium.com/go.php?id=544735>
2. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. Новосибирск: Изд-во Ин-та математики, 1999. 270 с.  
<https://lib.dvfu.ru:8443/lib/item?id=chamo:10172&theme=FEFU>
3. Клещев А.С. Математические основы информатики: Курс лекций. Находка: Институт технологии и бизнеса. – 2002. – 75 с.
4. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф.Хорошевский. СПб: Питер, 2001. 382 с.  
<https://lib.dvfu.ru:8443/lib/item?id=chamo:15439&theme=FEFU>
5. Кухаренко Б.Г. Интеллектуальные системы и технологии – М.: Московская государственная академия водного транспорта, 2015. – 116 с.  
<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 программное обеспечение: Microsoft Excel, Google Colaboratory, CPython.

## **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. Рекомендуется изучить структуру и основные положения Раб очей программы дисциплины. Обратите 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

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/зачету.** К сдаче экзамена/зачета допускаются обучающиеся, выполнившие все задания (лабораторные, 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 DLPPProjectorPT-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 Visial 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 Photoshopе 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**

Используются следующие оценочные средства:

Устный опрос:

1. Собеседование (УО-1)
2. Презентация / сообщение (УО-3)
3. Круглый стол (УО-4)

Письменные работы:

1. Лабораторная работа (ПР-6)
2. Тест (ПР-1)

### **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.

Презентация / сообщение (УО-3) – продукт самостоятельной работы обучающегося, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно-практической, учебно-исследовательской или научной темы.

Круглый стол (УО-4) - Оценочные средства, позволяющие включить обучающихся в процесс обсуждения спорного вопроса, проблемы и оценить их умение аргументировать собственную точку зрения.

### **Письменные работы**

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.

Тест (ПР-1) - Система стандартизированных заданий, позволяющая автоматизировать процедуру измерения уровня знаний и умений обучающегося.

### **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), зачет (4-й, весенний 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.
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 зачету

1. Индукция, дедукция, абдукция. Данные, знания, база данных, база знаний. Задача индуктивного формирования баз знаний.

2. Особенности, сходства и различия основных направлений исследований в области искусственного интеллекта (интеллектуальный анализ данных, обнаружение знаний в БД, машинное обучение, распознавание образов).
3. Этапы решения задач интеллектуального анализа данных. Примеры задач и областей применения методов интеллектуального анализа данных.
4. Модель предметной области, параметры модели, онтология. Прямая и обратная задачи математического моделирования. Внешние и внутренние критерии выбора моделей и признаков. Основные этапы моделирования.
5. Постановки задач классификации и диагностики. Сравнение постановок задач (по типу обучающих выборок, по результату работы, по объяснению результата работы). Правильная и точная классификация, типы ошибок: «ложная тревога» и «пропуск цели» (ошибки 1-ого и 2-ого рода).
6. Алгоритм обучения, решающее правило, обучающая и контрольная выборки. Свойства алгоритмов обучения. Проблема устойчивости. Постановки задачи проведения соревнований среди алгоритмов обучения (слабая и сильная).
7. Оценка внешних и внутренних свойств алгоритмов обучения. Связь оценок. Типы данных для экспериментов. Принципиальная схема компьютерного эксперимента для вычисления оценок свойств алгоритмов обучения.
8. Экспертный и индуктивный способы формирования баз знаний для интеллектуальных систем, их достоинства и недостатки. Проблема интерпретируемости результатов работы. Задачи подсистемы объяснения.
9. Постановка задачи медицинской диагностики. Серии экспериментов на модельных данных на примере решения задачи медицинской диагностики.
10. Способы визуального представления оценок внешних и внутренних свойств алгоритмов обучения в сериях экспериментов на модельных данных.

### **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, презентации, лабораторных работ, тестов) по оцениванию

фактических результатов обучения студентов.

Объектами оценивания выступают:

- степень усвоения теоретических знаний - оценивается в форме собеседования и тестирования;
- уровень овладения практическими умениями и навыками – оценивается в форме защиты лабораторных работ.

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.

### Примерные тесты для проверки сформированности компетенций

<p><b>. ОПК-7. Способен применять при решении профессиональных задач методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях.</b></p>	
<p>1. В принцип «Трёх V» не входит один из перечисленных признаков:</p>	<p>Ответы:</p> <p><b>а. Veracity (достоверность данных): в настоящее время достоверность имеющихся данных является важнейшим критерием для пользователей. Недостоверная информация приводит к затруднению анализа данных.</b></p> <p><b>б. Volume (объем):</b> накопленная база данных представляет собой гигантский объем информации, для которого обработка и хранение традиционными способами являются трудоёмкими процессами. Такой объем нуждается в новых подходах и в более усовершенствованных инструментах.</p> <p><b>в. Variety (многообразие):</b> данная характеристика означает возможность одновременной обработки структурированной и неструктурированной информации различных форматов. Главным отличием структурированной информации является возможность классификации.</p>
<p>2. К полуструктурированным данным относятся:</p>	<p>Ответы:</p> <p><b>а. Данные, которые не имеют определённой формы, могут включать в себя видео, аудио файлы, свободный текст, информацию, поступающую из</b></p>



	<p>социальных сетей.</p> <p><b>б. Данные, которые не соответствуют чёткой структуре таблиц и отношений в реляционных базах данных, однако такие данные содержат специальные теги и иные маркёры, позволяющие отделить семантические элементы.</b></p> <p>в. Данные, определяющие конкретную предметную область, упорядоченные специальным образом и организованные таким образом, чтобы над ними можно было выполнить анализ.</p>
--	---

<p><b>ПК-4. Владение навыками создания программного обеспечения для анализа и обработки информации.</b></p>	
<p>1. Предиктивное моделирование (Predictive Modelling) – это:</p>	<p><b>а. Процесс создания (или выбора) модели для предсказания вероятности наступления некоторого события.</b></p> <p>б. Компьютерная техника извлечения знаний, которая использует искусственный интеллект для распознавания образов и выделения значимых закономерностей из данных, находящихся в хранилищах или входных или выходных потоках.</p> <p>в. Методы и инструменты для перевода необработанной информации в осмысленную, удобную форму.</p>
<p>2. Text Mining – это:</p>	<p>Ответы:</p> <p><b>а. Нетривиальный процесс обнаружения действительно новых, потенциально полезных и понятных шаблонов в неструктурированных текстовых данных.</b></p> <p>б. Использование методов интеллектуального анализа для автоматического обнаружения веб-документов и услуг, извлечения информации из веб-ресурсов и сервисов.</p> <p>в. Собираемое название, используемое для обозначения совокупности методов обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.</p>

## Критерии оценки теста

Оценка	Требования
<i>отлично</i>	Более 80% правильных ответов
<i>хорошо</i>	65 – 80% правильных ответов
<i>удовл</i>	50 - 65% правильных ответов
<i>неудовл</i>	Менее 50% правильных ответов

100-86 баллов выставляется, если студент/группа точно определили содержание и составляющие части задания, умеют аргументированно отвечать на вопросы, связанные с заданием. Продемонстрировано знание и владение навыками самостоятельной исследовательской работы по теме. Фактических ошибок, связанных с пониманием проблемы, нет.

85-76 - баллов - работа студента/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы

60-50 баллов - если работа представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы

## Шкала оценивания

Менее 60 баллов	Незачтено	неудовлетворительно
От 61 до 75 баллов	Зачтено	удовлетворительно
От 76 до 85 баллов	Зачтено	хорошо
От 86 до 100 баллов	Зачтено	отлично