

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**  
федеральное государственное автономное образовательное учреждение  
высшего образования  
«Балтийский федеральный университет имени Иммануила Канта»  
Высшая школа компьютерных наук и искусственного интеллекта

**РАБОЧАЯ ПРОГРАММА МОДУЛЯ**

**Модуль 5. Искусственный интеллект**

**Шифр: 02.03.02**

**Направление подготовки: Фундаментальная информатика и информационные  
технологии**

**Профиль: Программная инженерия в искусственном интеллекте**

**Квалификация (степень) выпускника: бакалавр**

Калининград  
2023

## Лист согласования

### Составители:

- Верещагин Сергей Дмитриевич, к. ф.-м.н., доцент
- Верещагин Михаил Дмитриевич, к. ф.-м.н., доцент
- Мищук Богдан Ростиславович, к. ф.-м.н., доцент

Рабочая программа утверждена на заседании  
Ученого совета ОНК «Институт высоких технологий»

Протокол № 33 от «27» октября 2023 г.

Председатель Ученого совета ОНК «Институт высоких технологий»

Профессор, д.ф.-м.н.,  
руководитель ОНК «Институт высоких технологий»

А.В. Юров

Директор высшей школы компьютерных наук  
и искусственного интеллекта

М.Д. Верещагин

Руководитель ОПОП ВО

С.С. Головин

## Содержание

1. Название образовательного модуля «Модуль 5. Искусственный интеллект»
2. Характеристика образовательного модуля
3. Методические указания для обучающихся по освоению модуля
4. Программы дисциплин образовательного модуля
  - 4.1. Программа дисциплины «Методы машинного обучения»
  - 4.2. Программа дисциплины «Анализ и визуализация данных»
5. Программа практики
6. Программа итоговой аттестации по модулю

## 1. Название модуля: Модуль 5. Искусственный интеллект

### 2. Характеристика модуля

#### 2.1. Образовательные цели и задачи:

Модуль ставит своей целью создать условия для эффективного формирования и развития профессиональных компетенций.

Для достижения поставленной цели необходимо решить следующие задачи:

- Ознакомление студентов с комплексом современных технологий и концепций, в области прикладного искусственного интеллекта.
- Способствовать формированию у обучающихся практических навыков по статистическому анализу данных.
- Формировать у обучающихся базовые знания о современных задачах анализа данных и методов их решения, включая анализ соцсетей, текстов, построение ансамблей алгоритмов, в том числе с помощью алгебраического подхода к решению задач классификации.

#### 2.2. Образовательные результаты выпускника

Код компетенции	Результаты освоения образовательной программы (ИДК)	Результаты обучения по дисциплине
ПК-4. Способен разрабатывать и применять методы машинного обучения для решения задач	ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	<b>Знать:</b> основные принципы решения задач машинного обучения и анализа данных  <b>Уметь:</b> создавать алгоритмические модели типовых задач, проводить спецификацию задачи, реализовывать программы на алгоритмических языках высокого уровня, интерпретировать полученные результаты  <b>Владеть:</b> пониманием методов построения машинного обучения и анализа данных
ПК-7. Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-7.1. Осуществляет поиск данных в открытых источниках, специализированных библиотеках и репозиториях ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных	<b>Знать:</b> виды представления данных, методы поиска и парсинга данных, уровни представления данных (ODS, DDL, семантический слой, модель данных), основные инструменты, библиотеки и технологии Data Science <b>Уметь:</b> отделять достоверные источники данных от

	для машинного обучения	сомнительных, осуществлять критических отбор данных, проверять их на целостность и непротиворечивость, использовать инструменты и библиотеки для Data Science для поиска данных в открытых источниках, специализированных библиотеках и репозиториях <b>Владеть:</b> методами редукции размерности элементов набора данных и их предварительной статистической обработки, разметки структурированных и неструктурированных данных, а также методами планирования вычислительного эксперимента, формирования обучающей и контрольной выборок
--	------------------------	--

### 3. Методические указания для обучающихся по освоению модуля

Освоение дисциплин модуля закладывает базу для будущей профессиональной деятельности в сфере ИТ. Оно должно начинаться с внимательного ознакомления с рабочими программами дисциплин, обязательными компонентами которых являются: перечень тем, подлежащих усвоению; задания; списки учебных пособий и рекомендуемой литературы; списки контрольных вопросов, заданий.

При изучении дисциплин модуля необходимо последовательно переходить от дисциплины к дисциплине, от темы к теме, следуя внутренней логике, заложенной в программе дисциплины модуля. Только так можно достичь полного понимания материала, хорошей ориентации в специальной литературе, формирования собственной точки зрения и умений практического характера. Для более глубокого и эффективного освоения дисциплин рекомендуется предварительная подготовка к занятиям.

### 4. Программы дисциплин модуля

#### 4.1. Программа дисциплины «Методы машинного обучения».

##### 1. Наименование дисциплины: «Методы машинного обучения».

**Целью** курса «Методы машинного обучения» - дать слушателям широкий обзор задач и методов машинного обучения.

## 2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Код компетенции	Результаты освоения образовательной программы (ИДК)	Результаты обучения по дисциплине
ПК-4. Способен разрабатывать и применять методы машинного обучения для решения задач	ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	<b>Знать:</b> основные принципы решения задач машинного обучения и анализа данных  <b>Уметь:</b> создавать алгоритмические модели типовых задач, проводить спецификацию задачи, реализовывать программы на алгоритмических языках высокого уровня, интерпретировать полученные результаты  <b>Владеть:</b> пониманием методов построения машинного обучения и анализа данных

## 3. Место дисциплины в структуре образовательной программы

«Методы машинного обучения» представляет собой дисциплину обязательной части (Б1.О.05) направления подготовки бакалавриата 02.03.02 «Фундаментальная информатика и информационные технологии», профиль «Программная инженерия в искусственном интеллекте».

## 4. Виды учебной работы по дисциплине.

Виды учебной работы по дисциплине зафиксированы учебным планом основной профессиональной образовательной программы по указанному направлению и профилю, выражаются в академических часах. Часы контактной работы и самостоятельной работы обучающегося и часы, отводимые на процедуры контроля, могут различаться в учебных планах ОПОП по формам обучения. Объем контактной работы включает часы контактной аудиторной работы (лекции/практические занятия/ лабораторные работы), контактной внеаудиторной работы (контроль самостоятельной работы), часы контактной работы в период аттестации. Контактная работа, в том числе может проводиться посредством электронной информационно-образовательной среды университета с использованием ресурсов сети Интернет и дистанционных технологий

## 5. Содержание дисциплины, структурированное по темам (разделам)

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные

учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане). Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы.

5 семестр

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1	<b>Тема 1. Терминология</b>	Наука о данных (Data Science) Статистика (Statistics) Искусственный интеллект (Artificial Intelligence) Анализ данных (Data Mining) Машинное обучение (Machine learning) Большие данные (Big Data)
2	<b>Тема 2. Постановка основных задач</b>	Обучение с учителем (с размеченными данными / метками) целевая функция объект метка классификация Прогнозирование Пространство объектов признаков пространство Извлечение признаков Визуализация задач функции ошибки эмпирический риск обучающая выборка Задачи оптимизации в обучении Модель алгоритмов Алгоритм Обучение Обобщающая способность Схема решения задачи машинного обучения Как решаются задачи Обучение без учителя /с неразмеченными данными Обучение с частично размеченными данными трансдуктивное обучение Обучение с подкреплением Структурный вывод Активное обучение Онлайн-обучение

		<p>Transfer Learning  Multitask Learning  Feature Learning  Проблемы в машинном обучении  Примеры модельных задач</p>
3	<b>Тема 3.</b> Математика в машинном обучении: краткий обзор	<p>Бритва Оккама  Теорема о бесплатном сыре  Футбольный оракул  Сведения из ТВиМС  Задание распределений  Средние и отклонения  Условная плотность, маргинализация и обуславливание  Точечное оценивание  Оценка максимального правдоподобия  Дивергенция Кульбака-Лейблера ковариация и корреляция  Оценка плотности гистограммного подхода  Парзеновский подход  Нормальное распределение  Центральная предельная теорема  Теория информации  Проклятие размерности  Сингулярное разложение матрицы (SVD)  матричное дифференцирование</p>
4	<b>Тема 4.</b> Оптимизация	<p>Методы безусловной оптимизации  Методы нулевого порядка  Методы первого порядка  Методы второго порядка  Градиентный спуск  Наискорейший градиентный спуск  Стохастический градиентный спуск  Обучение: Пакетное, онлайн, по минибатчам  Метод градиентного спуска в машинном обучении  Стационарные точки  Метод Ньютона  Квази-ньютоновские методы  Оптимизация с ограничениями</p>
5	<b>Тема 5.</b> Метрические алгоритмы	<p>Метрические алгоритмы (distance-based)  Ближайший центроид (Nearest centroid algorithm)  Подход, основанный на близости  kNN в задаче классификации  kNN в задаче регрессии  Обоснование 1NN  Ленивые (Lazy) и нетерпеливые (Eager) алгоритмы  Весовые обобщения kNN  Различные метрики: Минковского, Евклидова, Манхэттенская, Махало-нобиса, Canberra distance, Хэмминга, косинусное,</p>

		<p>расстояние Джаккарда, DTW, Левенштейна</p> <p>Приложения метрического-го подхода:          нечёткий матчинг таблиц, Ленкор, в DL,          классификация текстов</p> <p>Эффективные методы поиска ближайших          соседей</p> <p>Регрессия Надарая-Ватсона</p>
6	<b>Тема 6. Линейные методы</b>	<p>Линейная регрессия</p> <p>Обобщённая линейная регрессия</p> <p>Проблема вырожденности матрицы</p> <p>Регуляризация. Основные виды          регуляризации</p> <p>Гребневая регрессия (Ridge Regression)</p> <p>LASSO (Least Absolute Selection and          Shrinkage Operator)</p> <p>Elastic Net</p> <p>Селекция признаков</p> <p>Ошибка с весами</p> <p>Устойчивая регрессия (Robust Regression)</p> <p>Линейные скоринговые модели в задаче          бинарной классификации</p> <p>Логистическая регрессия</p> <p>Probit-регрессия</p> <p>Многоклассовая логистическая регрессия</p> <p>Линейный классификатор</p> <p>Перцептрон</p> <p>Оценка функции ошибок через гладкую          функцию</p>

6 семестр

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1	<b>Тема 7. Деревья решений.</b>	<p>Деревья решений (CART)</p> <p>Предикаты / ветвления</p> <p>Ответы дерева</p> <p>Критерии расщепления в задачах          классификации: Missclassification criteria,          энтропийный, Джини</p> <p>Критерии остановки при построении          деревьев</p> <p>Проблема переобучения для деревьев</p> <p>Подрезка (post-pruning)</p> <p>Классические алгоритмы построения          деревьев решений: ID3, C5.0</p> <p>Важности признаков</p> <p>Проблема пропусков (Missing Values)</p> <p>Категориальные признаки</p> <p>Сравнение: деревья vs линейные модели</p>
2	<b>Тема 8. Контроль качества и выбор модели</b>	<p>Проблема контроля качества          выбора модели (Model Selection) в широком          смысле</p>

		<p>Правила разбиения вы-борки</p> <p>Отложенный контроль (held-out data, hold-out set)</p> <p>Скользящий контроль (cross-validation)</p> <p>Бутстреп (bootstrap)</p> <p>Контроль по времени (out-of-time-контроль)</p> <p>Локальный контроль</p> <p>Кривые обучения (Learning Curves)</p> <p>Перебор параметров</p>
3	<b>Тема 9.</b> Ансамблирование в машинном обучении	<p>Ансамбли алгоритмов: примеры и обоснование</p> <p>комитеты (голосование) / усреднение</p> <p>Бэггинг</p> <p>Кодировки / перекодировки ответов, ECOC</p> <p>Стекинг и блендинг</p> <p>Бустинг: AdaBoost, For-ward stagewise additive modeling (FSAM)</p> <p>«Ручные методы»</p> <p>Однородные ансамбли</p>
4	<b>Тема 10.</b> Методы, основанные на деревьях: случайный лес, бустинг	<p>Случайный лес, его параметры, их настройка</p> <p>Бэггинг и OOB (out of bag)</p> <p>Важность признаков</p> <p>Близость (Proximity) с помощью RF</p> <p>Extreme Random Trees</p> <p>Градиентный бустинг над деревьями, его параметры, современные реализации, Продвинутые методы оптимизации</p>
5	<b>Тема 11.</b> Введение в рекомендательные системы	<p>Рекомендательные системы</p> <p>Персонализация, онлайн и оффлайн рекомендации</p> <p>Рекомендация по контенту (content based methods)</p> <p>Коллаборативная фильтрация: GroupLens-алгоритм, SVD, SVD++, timeSVD++, адаптация SVD под социальные связи</p> <p>One-class recommendation</p> <p>Факторизационная машина, факторизационная машина с полями (FFM – field-aware factorization machine)</p> <p>Простые методы рекомендаций: FPM – Frequent Pattern Mining</p> <p>Deep Semantic Similarity Model (DSSM)</p> <p>Контекст рекомендации</p> <p>Knowledge-based Recommendations</p> <p>Важность объяснений (explanations)</p> <p>Использование дополнительной информации</p> <p>Современные тренды в практике построения рекомендательных систем</p>
6	<b>Тема 12.</b> Сложность алгоритмов, переобучение,	<p>Проблема обобщения</p> <p>Переобучение</p>

	смещение и разброс	Недообучение Сложность алгоритмов Смещение и разброс Способы борьбы с переобучением
--	--------------------	--

## 6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Рекомендуемая тематика учебных занятий *лекционного* типа (предусматривающих преимущественную передачу учебной информации преподавателями):

*5 семестр:*

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1	<b>Тема 1.</b> Терминология	Лекция 1. Понятийный аппарат
2	<b>Тема 2.</b> Постановка основных задач	Лекция 2. Обучение с учителем Лекция 3. Обучение без учителя Лекция 4. Обучение с подкреплением
3	<b>Тема 3.</b> Математика в машинном обучении: краткий обзор	Лекция 5. Бритва Оккама Лекция 6. Оценка максимального правдоподобия. Дивергенция Кульбака-Лейблера ковариация и корреляция Лекция 7. Теория информации
4	<b>Тема 4.</b> Оптимизация	Лекция 8. Методы безусловной оптимизации Лекция 9. Обучение: Пакетное, онлайн, по минибатчам Лекция 10. Оптимизация с ограничениями
5	<b>Тема 5.</b> Метрические алгоритмы	Лекция 11. Метрические алгоритмы (distance-based) Лекция 12. Ленивые (Lazy) и нетерпеливые (Eager) алгоритмы Лекция 13. Приложения метрического-го подхода: нечёткий матчинг таблиц, Ленкор, в DL, классификация текстов Лекция 14. Эффективные методы поиска ближайших соседей Регрессия Надарая-Ватсона
6	<b>Тема 6.</b> Линейные методы	Лекция 15. Линейная регрессия. Обобщённая линейная регрессия Лекция 16. Гребневая регрессия (Ridge Regression) Лекция 17. Устойчивая регрессия (Robust Regression)

		Лекция 18. Многоклассовая логистическая регрессия
--	--	---

*6 семестр*

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1	<b>Тема 7.</b> Деревья решений.	Лекция 1. Деревья решений (CART)
2	<b>Тема 8.</b> Контроль качества и выбор модели	Лекция 2. Проблема контроля качества Выбора модели (Model Selection) в широком смысле Лекция 3. Скользящий контроль (cross-validation) Бутстреп (bootstrap) Лекция 4. Кривые обучения (Learning Curves). Перебор параметров
3	<b>Тема 9.</b> Ансамблирование в машинном обучении	Лекция 5. Ансамбли алгоритмов: примеры и обоснование комитеты (голосование) / усреднение Бэггинг Лекция 6. Кодировки / перекодировки ответов, ECOC Стекинг и блендинг Лекция 7. Бустинг: AdaBoost, Forward stagewise additive modeling (FSAM) «Ручные методы» Однородные ансамбли
4	<b>Тема 10.</b> Методы, основанные на деревьях: случайный лес, бустинг	Лекция 8-9. Случайный лес, его параметры, их настройка Бэггинг и OOB (out of bag) Важность признаков Близость (Proximity) с помощью RF Extreme Random Trees Лекция 10-11. Градиентный бустинг над деревьями, его параметры, современные реализации, Продвинутое методы оптимизации
5	<b>Тема 11.</b> Введение в рекомендательные системы	Лекция 12. Рекомендательные системы Персонализация, онлайн и оффлайн рекомендации Рекомендация по контенту (content based methods) Лекция 13. Коллаборативная фильтрация: GroupLens-алгоритм, SVD, SVD++, timeSVD++, адаптация SVD под социальные связи One-class recommendation Факторизационная машина,

		факторизационная машина с полями (FFM – field-aware factorization machine) Лекция 14. Простые методы рекомендаций: FPM – Frequent Pattern Mining Deep Semantic Similarity Model (DSSM) Контекст рекомендации Knowledge-based Recommendations Важность объяснений (explanations) Использование дополнительной информации Лекция 15. Современные тренды в практике построения рекомендательных систем
6	<b>Тема 12.</b> Сложность алгоритмов, переобучение, смещение и разброс	Лекция 16-17. Проблема обобщения Переобучение Недообучение Лекция 18. Сложность алгоритмов Смещение и разброс Способы борьбы с переобучением

### **Рекомендуемая тематика *практических* занятий:**

Тематика практических занятий совпадает с темами лекционных занятий.

На практических занятиях решаются задачи по теме занятия.

### **Требования к самостоятельной работе обучающихся**

1. Работа с лекционным материалом, предусматривающая проработку конспекта лекций и учебной литературы, по всем темам из п. 6 настоящей рабочей программы.
2. Выполнение домашнего задания, предусматривающего решение задач, выполнение упражнений, выдаваемых на практических занятиях, по всем темам из п. 6 настоящей рабочей программы.

Руководствуясь положениями статьи 47 и статьи 48 Федерального закона от 29 декабря 2012 г. N 273-ФЗ «Об образовании в Российской Федерации» научно-педагогические работники и иные лица, привлекаемые университетом к реализации данной образовательной программы, пользуются предоставленными академическими правами и свободами в части свободы преподавания, свободы от вмешательства в профессиональную деятельность; свободы выбора и использования педагогически обоснованных форм, средств, методов обучения и воспитания; права на творческую инициативу, разработку и применение авторских программ и методов обучения и воспитания в пределах реализуемой образовательной программы и отдельной дисциплины.

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия,

практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане).

Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы.

## **7. Методические рекомендации по видам занятий**

Лекционные занятия.

В ходе лекционных занятий обучающимся рекомендуется выполнять следующие действия. Вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации по их применению. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Желательно оставить в рабочих конспектах поля, на которых во внеаудиторное время можно сделать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений.

Практические и семинарские занятия.

На практических и семинарских занятиях в зависимости от темы занятия выполняется поиск информации по решению проблем, практические упражнения, контрольные работы, выработка индивидуальных или групповых решений, итоговое обсуждение с обменом знаниями, участие в круглых столах, разбор конкретных ситуаций, командная работа, представление портфолио и т.п.

Самостоятельная работа.

Самостоятельная работа осуществляется в виде изучения литературы, эмпирических данных по публикациям и конкретным ситуациям из практики, подготовке индивидуальных работ, работа с лекционным материалом, самостоятельное изучение отдельных тем дисциплины; поиск и обзор литературы и электронных источников; чтение и изучение учебника и учебных пособий.

## **8. Фонд оценочных средств**

### **8.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы в рамках учебной дисциплины**

Основными этапами формирования указанных компетенций при изучении обучающимися дисциплины являются последовательное изучение содержательно

связанных между собой тем учебных занятий. Изучение каждой темы предполагает овладение обучающимися необходимыми компетенциями. Результат аттестации обучающихся на различных этапах формирования компетенций показывает уровень освоения компетенций.

Контролируемые разделы (темы) дисциплины	Индекс контролируемой компетенции (или её части)	Оценочные средства по этапам формирования компетенций
		текущий контроль по дисциплине
<b>Тема 1.</b> Терминология	ПК-4	тест
<b>Тема 2.</b> Постановка основных задач	ПК-4	тест
<b>Тема 3.</b> Математика в машинном обучении: краткий обзор	ПК-4	тест
<b>Тема 4.</b> Оптимизация	ПК-4	тест
<b>Тема 5.</b> Метрические алгоритмы	ПК-4	тест
<b>Тема 6.</b> Линейные методы	ПК-4	тест
<b>Тема 7.</b> Деревья решений.	ПК-4	тест
<b>Тема 8.</b> Контроль качества и выбор модели	ПК-4	тест
<b>Тема 9.</b> Ансамблирование в машинном обучении	ПК-4	тест
<b>Тема 10.</b> Методы, основанные на деревьях: случайный лес, бустинг	ПК-4	тест
<b>Тема 11.</b> Введение в рекомендательные системы	ПК-4	тест
<b>Тема 12.</b> Сложность алгоритмов, переобучение, смещение и разброс	ПК-4	тест

## 8.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности процессе текущего контроля

### Примерные тестовые задания:

#### 5 семестр:

- Случайная величина принимает значение из отрезка  $[0,1]$ , её плотность линейная функция на этом отрезке, в нуле обращается в ноль. Чему равно матожидание с.в.? Совет: здесь и ниже, кроме аналитического решения напишите на Python программу для оценки названных параметров.

0

1/2

\* 2/3

3/4

4/5

1

- Чему равна мода этой с.в.?

0

1/2

2/3

3/4

4/5

\* 1

нет правильного ответа

- Чему равна медиана этой с.в.?

0

1/2

2/3

3/4

4/5

1

\* нет правильного ответа

- Чему равна дисперсия этой с.в.?

1/18

- Предположим, что в задаче бинарной классификации с одним признаком объекты класса 1 распределены так, как описано выше, а объекты класса 0 распределены равномерно на отрезке  $[0,1]$ . Оба класса равновероятны. Какой оптимальный порог для отнесения объектов к классу 1 (выше него считаем, что они из класса 1), если оба класса равновероятны?

1/3

\* 1/2

2/3

3/4

4/5

- Что лучше использовать для определения монотонной зависимости между переменными?

Корреляционный коэффициент Пирсона

\* Коэффициент корреляции Спирмена

оценку ММП (MLE)

- Парзенковский подход используется для параметрического оценивания плотности
- \* непараметрического оценивания плотности  
оценки плотности смеси распределений

- Чему равна оценка плотности в точке 1.5 для выборки {1, 2, 3} парзеновским методом с треугольным ядром (радиус основания ядра 1)?

0

1/6

\* 1/3

2/3

1

нет правильного ответа

- Пусть случайная величина равна сумме двух равномерно распределённых величин на отрезке [0,1]. Как выглядит её плотность распределения?

«колокольчик»

\* треугольник

трапеция

прямоугольник

10. К чему стремится угол между соседними диагоналями n-мерного гиперкуба при увеличении размерности?

\* 0

$\pi/4$

$\pi/2$

нет правильного ответа

- Запишите сумму квадратов сингулярных чисел для матрицы  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ .

2

- При минимизации функции  $x^2$  методом градиентного спуска с темпом 1.0 и начальной точкой 1.0, какая будет оценка  $\text{argmin}$  после 4й итерации?

1

- Выберите верные фразы:

+ в SGD случайный порядок объектов

+ SGD может использоваться при онлайн-оптимизации (обучении)

+ SGD может использоваться для минимизации суммы ошибок на объектах моделей классификации / регрессии

SGD – метод оптимизации второго порядка

SGD – это метод условной оптимизации

13. Решите задачу условной оптимизации:  $x^2 + y^2 \rightarrow \min$  при условии  $x+2y \geq 1$  (ответ - значение функции в точке минимума).

0.2

- Дана обучающая выборка в однопризнаковой задаче бинарной классификации. Объекты первого класса – {0, 4, 4.5}, второго – {3, 7, 8}. Какая оценка точности алгоритма 1NN методом LOO(leave-one-out)? Ответ умножьте на 6.

4

- В предыдущей задаче оцените точность алгоритма ближайшего центроида.

4

16. Пусть в однородной области метрического пространства объектов бинарной задачи классификации вероятность того, что случайно выбранный объект принадлежит классу 1, равна 1/4. Чему равна ошибка алгоритма ближайшего соседа в этой области?

0.375

17. Что такое ленивый алгоритм (Lazy learner)?

- медленно классифицирует

+ не формирует модель описания данных во время обучения

- синоним для алгоритма kNN

- который по умолчанию не использует метод предсказания меток (predict)

18. Задачу регрессии можно решать с помощью (выбрать все верные варианты)

+ kNN

+ метода Надарая-Ватсона

- метода логистической регрессии

19. Какие расстояния численно наибольшие для пары точек (1,1) и (2,2):

- Евклидово

- Чебышева

+ Манхэттенское

20. Пусть даны векторы (1,1,2,2,3,3), (1, 4). Чему равно расстояние DTW?

4

21. Решите матричное уравнение  $Xw=y$ ,  $W=[[1, 1], [1, 2], [1, 3]]$ ,  $y=[1,2,1]$  с помощью минимизации невязки. В ответ запишите скалярное произведение вектора  $w$  и вектора  $(3, -1)$ .

4

22. В регуляризации по Тихонову:

- + к оптимизируемому функционалу добавляется специальное слагаемое
- вводится ограничение на норму вектора-решения
- оптимизируемый функционал оценивается сверху

23. В гребневой регрессии (выберите верные утверждения)

- + Используется регуляризация по Тихонову
- + матрица Грамма в псевдообратной становится невырожденной
- происходит гарантированное зануление элементов вектора решения
- происходит гарантированная селекция признаков

24. С помощью персептронного алгоритма решите систему уравнений  $a+b>0$ ,  $3a-b>0$ ,  $a-b<0$ . Начальное приближение  $(a,b) = (0,0)$ , неравенства просматриваются слева направо. В ответ запишите значение  $b/a$ .

2

25. Выберите верные фразы:

- + Для селекции признаков можно использовать LASSO
- + Устойчивая регрессия (Robust Regression) хороша в задаче с выбросами
- В логистической регрессии минимизируют среднее квадратичное отклонение ответов модели от истинных меток
- решение линейной регрессии робастно (устойчиво к выбросам)

## 6 семестр

1. Что особенного в деревьях решений вида «oblique decision trees»?

- ограничение на глубину
- использование предварительной обрезки (pre-pruning)
- + специальный предикат ветвления
- возможность распараллеливания при построении

2. Почему при построении дерева используют рекурсивную жадную стратегию?

- алгоритма оптимизации не существует

+ задача построения оптимального дерева очень сложна (в одном частном случае это NP-полная проблема)

- это, как правило, быстрее градиентного спуска

3. Рассмотрим 10 объектов, если их упорядочить по первому признаку, то их метки будут чередоваться следующим образом: [0,0,1,0,1,0,1,1,1,0]. Найдите максимальное значение критерия расщепления Missclassification criteria.

0.2

4. В предыдущей задаче найдите максимальное значение критерия расщепления Gini (основанного на мере неоднородности Gini).

0.125

5. В предыдущей задаче найдите максимальное значение энтропийного критерия расщепления. Ответ округлите до первой цифры после запятой (например, 0.1).

0.2

- Отметьте верные утверждения:

+ Обрезку (post-pruning) используют крайне редко

+ Деревья особенно эффективны в ансамбле

+ Деревья – нестабильный (неустойчивый) алгоритм

- Деревья часто используют для экстраполяции

- Что такое C5.0?

+ алгоритм построения деревьев решений

- специальный критерий расщепления

- запатентованный способ обрезки деревьев

- метод регуляризации при построении деревьев

8. Пусть есть категориальный признак со значениями [A, A, B, B, C, C, D, D], с целевыми значениями [2, 0, 2, 8, 3, 5, 0, 4]. Какое будет расщепление со стандартным критерием, использующим дисперсию?

- A, B | C, D

- A, C | B, D

+ A, D | B, C

- A | C, B, D

- A, B, C | D

- Что из перечисленного можно использовать для выбора модели (Model Selection):

+ бутстреп

- регуляризацию
- + разбиение на фолды

10. В какой из перечисленных функций библиотеки scikit-learn схема контроля гарантирует определённую пропорцию объектов разных классов?

- ShuffleSplit
- GroupShuffleSplit
- + StratifiedShuffleSplit
- KFold
- GroupKFold
- + StratifiedKFold
- PredefinedSplit

11. Пусть по транзакциям пользователя мы предсказываем его покупательскую активность в следующем методе. Какие способы контроля следует выбрать (в данных – статистика по всем клиентам банка)?

- LOOCV
- + out of time
- + out of sample

12. Кривые обучения (Learning Curves) могут

- + оценить достаточность объёма выборки
- + оценить, не слишком ли простая модель использована
- подобрать оптимальные значения всех параметров
- + оценить переобучение алгоритма

13. Пусть дана выборка целевых значений: 1, 3, 2, 1 (упорядочено по времени получения меток). Используется модель константных алгоритмов (ответ равен среднему по всем меткам обучения). Функция ошибки – MAE (средний модуль отклонения). Чему равна средняя ошибка при контроле по времени (Out-of-time)?

1

- В предыдущей задаче – чему будет равна ошибка LOOCV (контроля по одному)?

1

15. Чему равна ошибка комитета большинства над тремя алгоритмами бинарной классификации с вероятностями ошибки 0.3, 0.2, 0.1 (действует сильное предположение о независимости ответов алгоритмов)?

0.0098

16. В какой модели разнообразие базовых алгоритмов повышается за счёт варьирования обучающей выборки?

- + бэггинг (Bagging)
- + метод случайных подпространств (Random Subspaces)
- + случайные леса (Random Forests)
- нейросети

17. В какой модели производится перекодировка целевого признака?

- комитеты
- + ЕСОС
- стекинг
- бустинг

18. В какой модели применяется взятие бустреп-подвыборок?

- + бэггинг (Bagging)
- + случайные леса (RF)
- стекинг
- Feature-Weighted Linear Stacking

19. Для чего можно использовать ООВ-предсказания?

- + для оценки качества модели
- + для реализации стекинга
- для регуляризации
- для кодирования целевого вектора
- + для вычисления рейтинга (важности) признаков

20. Какие из перечисленных ниже моделей являются последовательными ансамблями (Sequential ensembles)?

- бэггинг (Bagging)
- + Adaboost
- + градиентный бустинг
- нейронные сети
- случайные леса (RF)

21. Пусть случайные величины одинаково распределены (среднее равно 1, дисперсия – 2), корреляция между любой парой величин равна 0.1. К чему стремится среднее арифметическое этих величин при увеличении числа наблюдений (т.е. увеличении числа этих величин)?

0.2

22. В каких моделях увеличение числа базовых алгоритмов не приводит к переобучению?

- стекинг
- бустинг
- + случайные леса

23. Что происходит при увеличении глубины деревьев (считаем, что в ансамблях достаточное число деревьев)?

- + как правило, увеличивается качество случайного леса на тесте
- + как правило, увеличивается качество случайного леса на обучении
- как правило, увеличивается качество бустинга над деревьями на тесте
- + как правило, увеличивается качество бустинга на обучении

24. В одном из подходов к оценке важности признаков используют перестановку значений. Почему именно перестановку (а, например, не анализ качества алгоритмов без соответствующего признака)?

- + это не меняет распределение по признаку
- + это позволяет не переучивать модель
- это гарантирует стабильность модели
- это гарантирует такое же распределение ответов модели

25. Чем экстремальные леса (Extreme Random Trees) отличаются от случайных (Random Forest)?

- не используем критерии расщепления (типа gini и энтропийного)
- + быстрее построение ансамбля
- нужен градиент функции ошибки
- + качество, как правило, чуть хуже

26. Что используется в продвинутых методах реализации градиентного бустинга (как в библиотеке XGBoost)?

- + принцип минимальной длины (MDL)
- + вторые производные функции ошибки
- автоматический выбор ключевых параметров, например learning\_rate

27. На какие слагаемые раскладывается квадратичная ошибка регрессора (матожидание квадрата разности прогноза и истинного значения)?

- + шум (noise)
- квадрат шума
- + разброс (variance)
- квадрат разброса
- смещение (bias)

+ квадрат смещения

28. При повышении числа соседей к метода kNN...

- увеличивается сложность модели
- + увеличивается качество на обучении
- увеличивается качество на контроле
- увеличивается разброс (variance)
- + увеличивается стабильность

29. Что из перечисленного приводит к уменьшению переобучения?

- + аугментация
- + регуляризация
- + увеличение объёма выборки

### Пример практического задания

Задача машинного обучения с реальными данными, выложенная на

<https://inclass.kaggle.com/c/dayofweek/>

#### *Описание*

Для 300000 пользователей дана статистика посещений ресурса за 1099 дней. Необходимо предсказать день недели следующего визита.

#### *Метрика качества*

Используется простой процент правильных ответов. Например,  
 $\text{performance}([1,2,2,7], [3,2,2,7]) = 0.75$

#### *Формат ответа*

В загружаемом файле по строкам перечислены идентификаторы пользователей и номера дней их первых визитов по версии вашего алгоритма:

id,nextvisit

1, 7

#### *Данные*

В файле train.csv перечислены даты визитов пользователей. Каждая строка - информация по одному пользователю. Сначала идёт id, потом через пробел номера дней, когда были визиты. Нумерация идёт от некоторого фиксированного момента. Номера могут быть от 1 до 1099 (т.е. статистика охватывает период примерно 3 года). Первый день в нумерации - понедельник.

Необходимо предсказать день недели первого визита после 1099го дня, т.е. для каждого пользователя вычислить

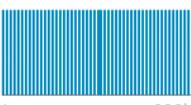
0 - нет визита

1 - понедельник

- 2 - вторник
- 3 - среда
- 4 - четверг
- 5 - пятница
- 6 - суббота
- 7 - воскресенье

### Данные в просмотрщике платформы inclass.kaggle

train.csv (142.76 MB)

	id	visits
		300000 unique values
1	1	30 84 126 135 137 179 242 342 426 456 460 462 483 594 600 604 704 723 744 787 804 886 924 928 946 954 1039 1040 1052
2	2	24 53 75 134 158 192 194 211 213 238 251 305 404 418 458 476 493 571 619 731 739 759 761 847 883 943 962 981 983 1036 1046
3	3	51 143 173 257 446 491 504 510 559 616 719 735 769 800 833 853 856 867 882 916 929 937 944 954 956 968 1007 1052 1071 1078 1094 1096

### 8.3. Перечень вопросов и заданий для промежуточной аттестации по дисциплине

5 семестр

Промежуточная аттестация осуществляется в форме зачет

В качестве средств, используемых на промежуточной аттестации предусматривается:

Билеты

6 семестр

Промежуточная аттестация осуществляется в форме экзамена

В качестве средств, используемых на промежуточной аттестации предусматривается:

Билеты

**5 семестр**

## Вопросы к зачету:

- Терминология: Наука о данных (Data Science), Статистика (Statistics), Искусственный интеллект (Artificial Intelligence), Анализ данных (Data Mining), Машинное обучение (Machine learning), Большие данные (Big Data)
- Обучение с учителем (с размеченными данными / метками): целевая функция, объект, метка, классификация, прогнозирование
- Пространство объектов, признаковое пространство, извлечение признаков, визуализация задач
- Функции ошибки, эмпирический риск, обучающая выборка, задачи оптимизации в обучении, обобщающая способность
- Модель алгоритмов, алгоритм, обучение, схема решения задачи машинного обучения
- Обучение без учителя / с неразмеченными данными, обучение с частично размеченными данными, трансдуктивное обучение
- Обучение с подкреплением, структурный вывод, активное обучение, онлайн-обучение, Transfer Learning, Multitask Learning, Feature Learning
- Математика в машинном обучении: бритва Оккама, теорема о бесплатном сыре, футбольный оракул, теория информации, проклятие размерности, сингулярное разложение матрицы (SVD), матричное дифференцирование
- Сведения из ТВиМС: задание распределений, средние и отклонения, условная плотность, маргинализация и обуславливание, точечное оценивание, оценка максимального правдоподобия, дивергенция Кульбака-Лейблера, ковариация и корреляция, нормальное распределение, центральная предельная теорема
- Оценка плотности: гистограммный подход, Парзеновский подход
- Оптимизация: методы безусловной оптимизации, нулевого порядка, первого порядка, второго порядка, метод градиентного спуска в машинном обучении, стационарные точки, метод Ньютона, квази-ньютоновские методы, оптимизация с ограничениями
- Градиентный спуск, наискорейший градиентный спуск, стохастический градиентный спуск, обучение: Пакетное, онлайн, по минибатчам
- Метрические алгоритмы (distance-based), ближайший центроид (Nearest centroid algorithm), подход, основанный на близости, kNN в задаче классификации / регрессии, обоснование 1NN, ленивые (Lazy) и нетерпеливые (Eager) алгоритмы
- Весовые обобщения kNN, регрессия Надарая-Ватсона
- Различные метрики: Минковского, Евклидова, Манхэттенская, Махаланобиса, Canberra distance, Хэмминга, косинусное, расстояние Джаккарда, DTW, Левенштейна, приложения метрического подхода: нечёткий матчинг таблиц, Ленкор, в DL, классификация текстов, эффективные методы поиска ближайших соседей

- Линейные методы: линейная регрессия, обобщённая линейная регрессия, проблема вырожденности матрицы, регуляризация, основные виды регуляризации, гребневая регрессия (Ridge Regression), LASSO (Least Absolute Selection and Shrinkage Operator), Elastic Net
- Селекция признаков, ошибка с весами, устойчивая регрессия (Robust Regression)
- Линейные скоринговые модели в задаче бинарной классификации, логистическая регрессия, Probit-регрессия, многоклассовая логистическая регрессия
- Линейный классификатор, перцептрон, оценка функции ошибок через гладкую функцию

## 6 семестр

### Вопросы к экзамену

- Деревья решений (CART), предикаты / ветвления, ответы дерева, критерии расщепления в задачах классификации: Missclassification criteria, энтропийный, Джини, критерии остановки при построении деревьев, проблема переобучения для деревьев, подрезка (post-pruning), классические алгоритмы построения деревьев решений: ID3, C5.0
- Важности признаков, проблема пропусков (Missing Values), категориальные признаки, сравнение: деревья vs линейные модели
- Проблема контроля качества, выбора модели (Model Selection) в широком смысле, правила разбиения выборки, кривые обучения (Learning Curves)
- перебор параметров
- Отложенный контроль (held-out data, hold-out set), скользящий контроль (cross-validation), бутстреп (bootstrap), контроль по времени (out-of-time-контроль), локальный контроль
- Ансамбли алгоритмов: примеры и обоснование, комитеты (голосование) / усреднение, бэггинг, кодировки / перекодировки ответов, ECOC
- Стекинг и блендинг, бустинг: AdaBoost, Forward stagewise additive modeling (FSAM), «Ручные методы», однородные ансамбли
- Случайный лес, его параметры, их настройка, бэггинг и OOB (out of bag), важность признаков, близость (Proximity) с помощью RF, Extreme Random Trees
- Градиентный бустинг над деревьями, его параметры, современные реализации, продвинутые методы оптимизации
- Рекомендательные системы, персонализация, онлайн и оффлайн рекомендации, рекомендация по контенту (content based methods), One-class recommendation, использование дополнительной информации, современные тренды в практике построения рекомендательных систем
- Коллаборативная фильтрация: GroupLens-алгоритм, SVD, SVD++, timeSVD++, адаптация SVD под социальные связи

- Факторизационная машина, факторизационная машина с полями (FFM – field-aware factorization machine)
- Простые методы рекомендаций: FPM – Frequent Pattern Mining, Deep Semantic Similarity Model (DSSM), контекст рекомендации, Knowledge-based Recommendations, важность объяснений (explanations)
- Сложность алгоритмов, переобучение, смещение и разброс: проблема обобщения, переобучение, недообучение, сложность алгоритмов, смещение и разброс, способы борьбы с переобучением

### Задачи к экзамену

Пусть случайная величина равна сумме двух независимых равномерно распределённых величин, одна – на отрезке  $[0,1]$ , вторая – на отрезке  $[0,2]$ . Как выглядит её плотность распределения?

«колокольчик»	треугольник	трапеция	прямоугольник
---------------	-------------	----------	---------------

В предыдущей задаче пусть указанные распределения – распределения классов 0 и 1 в задаче бинарной классификации. Оба класса равновероятны. Какая вероятность, что объект  $x=1$  принадлежит классу 0?

1/2	2/3	3/4	1
-----	-----	-----	---

При минимизации функции  $x^2$  методом градиентного спуска с темпом 0.5 и начальной точкой 1.0, какая будет оценка  $\text{argmin}$  после 1й итерации?

- 0.5	0	0.5	1
-------	---	-----	---

Выберите верные фразы

Для селекции признаков обычно используют L2-регуляризацию	Логистическая регрессия – ленивый алгоритм
Евклидово расстояние – частный случай расстояния Махаланобиса	С помощью перцептронного алгоритма можно решать системы линейных уравнений

Чему рано максимальное значение MC (Missclassification criteria)?

0	0.5	e	1
---	-----	---	---

Пусть дана выборка целевых значений: 1, 2, 3 (упорядочено по времени получения меток). Используется модель константных алгоритмов (ответ равен среднему по всем меткам обучения). Функция ошибки – MAE (средний модуль отклонения). Чему равна средняя ошибка при контроле LOOCV (контроля по одному)?

0.5	2/3	1	3/2
-----	-----	---	-----

В каком ансамбле следует использовать неустойчивые модели?

бэггинг	случайные леса	бустинг	ЕСОС
---------	----------------	---------	------

Что происходит при увеличении числа деревьев в градиентном бустинге (отметьте все варианты)?

ошибка на обучении падает	ошибка на контроле падает	ошибка на обучении возрастает	ошибка на контроле возрастает
---------------------------	---------------------------	-------------------------------	-------------------------------

Выберите верные фразы:

Критерий gini используется для построения деревьев в задаче регрессии	В экстремальных лесах (Extreme Random Trees) используется вычисление градиента ошибки
Контроль по фолдам используется для отбора модели	Аугментация – способ увеличения обучающей выборки

Пример экзаменационного билета

- Ансамбли алгоритмов: примеры и обоснование, комитеты (голосование) / усреднение, бэггинг, кодировки / перекодировки ответов, ЕСОС
- Стекинг и блендинг, бустинг: AdaBoost, Forward stagewise additive modeling (FSAM), «Ручные методы», однородные ансамбли
- Пусть дана выборка целевых значений: 1, 2, 3 (упорядочено по времени получения меток). Используется модель константных алгоритмов (ответ равен среднему по всем меткам обучения). Функция ошибки – MAE (средний модуль отклонения). Чему равна средняя ошибка при контроле LOOCV (контроля по одному)?

0.5	2/3	1	3/2
-----	-----	---	-----

#### 8.4. Планируемые уровни сформированности компетенций обучающихся и критерии оценивания

Уровни	Содержательное описание уровня	Основные признаки выделения уровня (этапы формирования компетенции, критерии оценки сформированности)	Пятибалльная шкала (академическая) оценка	Двухбалльная шкала, зачет	БРС, % освоения (рейтинговая оценка)
Повышенный	Творческая деятельность	<i>Включает нижестоящий уровень.</i> Умение самостоятельно принимать решение, решать проблему/задачу теоретического и прикладного характера на основе изученных методов, приемов, технологий	отлично	зачтено	86-100
Базовый	Применение знаний и умений в более широких контекстах учебной и профессиональной деятельности, нежели по образцу с большей степени самостоятельности и инициативы	<i>Включает нижестоящий уровень.</i> Способность собирать, систематизировать, анализировать и грамотно использовать информацию из самостоятельно найденных теоретических источников и иллюстрировать ими теоретические положения или обосновывать практику применения	хорошо		71-85
Удовлетворительный (достаточный)	Репродуктивная деятельность	Изложение в пределах задач курса теоретически и практически контролируемого материала	удовлетворительно		55-70
Недостаточный	Отсутствие признаков удовлетворительного уровня		неудовлетворительно	не зачтено	Менее 55

## 9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.

### Основная литература

- Авдеенко, Т. В. Введение в искусственный интеллект и логическое программирование. Программирование в среде Visual Prolog : учебное пособие / Т. В. Авдеенко, М. Ю. Целебровская. - Новосибирск : Изд-во НГТУ, 2020. - 64 с. - ISBN 978-5-7782-4182-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1869259> (дата обращения: 04.04.2023). – Режим доступа: по подписке.
- Веревкин, А. П. Искусственный интеллект в задачах моделирования, управления, диагностики технологических процессов : монография / А. П. Веревкин, Т. М. Муртазин. - Москва ; Вологда : Инфра-Инженерия, 2023. - 232 с. - ISBN 978-5-9729-1428-9. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2094393> (дата обращения: 18.11.2023). – Режим доступа: по подписке.

### Дополнительная литература

- Пенькова, Т. Г. Модели и методы искусственного интеллекта : учебное пособие / Т. Г. Пенькова, Ю. В. Вайнштейн. - Красноярск : Сиб. федер. ун-т, 2019. - 116 с. - ISBN 978-5-7638-4043-8. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1816605> (дата обращения: 04.04.2023). – Режим доступа: по подписке.
- 2.Сергеев, Н. Е. Системы искусственного интеллекта. Часть 1: Учебное пособие / Сергеев Н.Е. - Таганрог:Южный федеральный университет, 2016. - 118 с.: ISBN 978-5-9275-2113-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/991954> (дата обращения: 04.04.2023). – Режим доступа: по подписке.
- Сопов, Е. А. Многокритериальные нейроэволюционные системы в задачах машинного обучения и человеко-машинного взаимодействия : монография / Е. А. Сопов, И. А. Иванов. - Красноярск : Сиб. федер. ун-т, 2019. - 160 с. - ISBN 978-5-7638-3969-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1818898> (дата обращения: 04.04.2023). – Режим доступа: по подписке.

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

- НЭБ Национальная электронная библиотека, диссертации и прочие издания
- eLIBRARY.RU Научная электронная библиотека, книги, статьи, тезисы докладовконференций
- Гребенников Электронная библиотека ИД журналы
- ЭБС Консультант студента
- ПРОСПЕКТ ЭБС
- ЭБС ZNANIUM.COM
- ЭБС IBOOKS.RU

- РГБ Информационное обслуживание по МБА
- БЕН РАН
- Электронно-библиотечная система (ЭБС) Кантиана (<https://elib.kantiana.ru/>)

## **11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине.**

Программное обеспечение обучения включает в себя:

- система электронного образовательного контента БФУ им. И. Канта – [www.lms.kantiana.ru](http://www.lms.kantiana.ru), обеспечивающую разработку и комплексное использование электронных образовательных ресурсов;
- серверное программное обеспечение, необходимое для функционирования сервера и связи с системой электронного обучения через Интернет;
- корпоративная платформа;
- установленное на рабочих местах обучающихся ПО: Microsoft Windows 7, Microsoft Office Standart 2010, антивирусное программное обеспечение Kaspersky Endpoint Security.
- СУБД PostgreSQL (Свободное ПО, лицензия - Freeware).
- MongoDB (Свободное ПО, лицензия - Freeware).
- Python 2.7.15 (Anaconda2 5.2.0 64-bit)
- Python 3.6.5 (Anaconda3 5.2.0 64-bit)

## **12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.**

Для проведения занятий лекционного типа, практических и семинарских занятий используются специальные помещения (учебные аудитории), оборудованные техническими средствами обучения – мультимедийной проекционной техникой. Для проведения занятий лекционного типа используются наборы демонстрационного оборудования.

Для проведения лабораторных работ, (практических занятий – при необходимости) используются специальные помещения (учебные аудитории), оснащенные специализированным лабораторным оборудованием: персональными компьютерами с возможностью выхода в интернет и с установленным программным обеспечением, заявленным в п.11.

Для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используются специальные помещения (учебные аудитории), оборудованные специализированной мебелью (для обучающихся), меловой / маркерной доской.

Для организации самостоятельной работы обучающимся предоставляются помещения, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.

Для обучения инвалидов и лиц с ограниченными возможностями здоровья университетом могут быть представлены специализированные средства обучения, в том числе технические средства коллективного и индивидуального пользования.

## 4.2. Программа дисциплины «Анализ и визуализация данных»

### 1. Наименование дисциплины: «Анализ и визуализация данных».

**Целью** курса формирование у студентов теоретических знаний о принципах, методах, инструментах эффективной подготовки, анализа и визуализации данных, разработки дашбордов для решения задач управления, практических навыков анализа данных и разработки дашбордов в современных программных средах при решении профессиональных бизнес-задач.

### 2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Код компетенции	Результаты освоения образовательной программы (ИДК)	Результаты обучения по дисциплине
ПК-7. Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-7.1. Осуществляет поиск данных в открытых источниках, специализированных библиотеках и репозиториях ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	<b>Знать:</b> виды представления данных, методы поиска и парсинга данных, уровни представления данных (ODS, DDL, семантический слой, модель данных), основные инструменты, библиотеки и технологии Data Science <b>Уметь:</b> отделять достоверные источники данных от сомнительных, осуществлять критических отбор данных, проверять их на целостность и непротиворечивость, использовать инструменты и библиотеки для Data Science для поиска данных в открытых источниках, специализированных библиотеках и репозиториях <b>Владеть:</b> методами редукции размерности элементов набора данных и их предварительной статистической обработки, разметки структурированных и неструктурированных данных, а также методами планирования вычислительного эксперимента, формирования обучающей и контрольной выборки

### 3. Место дисциплины в структуре образовательной программы

«Анализ и визуализация данных» представляет собой дисциплину обязательной части (Б1.О.05) направления подготовки бакалавриата 02.03.02 «Фундаментальная информатика и информационные технологии», профиль «Программная инженерия в искусственном интеллекте».

#### **4. Виды учебной работы по дисциплине.**

Виды учебной работы по дисциплине зафиксированы учебным планом основной профессиональной образовательной программы по указанному направлению и профилю, выражаются в академических часах. Часы контактной работы и самостоятельной работы обучающегося и часы, отводимые на процедуры контроля, могут различаться в учебных планах ОПОП по формам обучения. Объем контактной работы включает часы контактной аудиторной работы (лекции/практические занятия/ лабораторные работы), контактной внеаудиторной работы (контроль самостоятельной работы), часы контактной работы в период аттестации. Контактная работа, в том числе может проводиться посредством электронной информационно-образовательной среды университета с использованием ресурсов сети Интернет и дистанционных технологий

#### **5. Содержание дисциплины, структурированное по темам (разделам)**

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане). Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы.

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
-------	--	-------------------------------------

•	1 Основные понятия интеллектуального анализа данных	Классификация задач интеллектуального анализа данных. Расстояния, сходства, метрики для различных типов данных. Этапы подготовки данных. Выдвижение гипотез. Методы сбора и систематизации фактов. Методы проведения экспертиз для выявления наиболее значимых факторов. Понятия парциальной и комплексной обработки. Анализ качества полученных моделей.
•	2 Использование NumPy и Pandas для анализа данных	Введение в NumPy. Векторизованные вычисления. Агрегирующие функции. Broadcasting. Логические операции. Введение в pandas. Индексы в pandas. Иерархические (множественные) индексы. Иерархические колонки. Агрегирование в pandas. Объединение нескольких наборов данных. Решение базовых задач анализа данных.
•	3 Кластерный анализ	Эвристические алгоритмы кластерного анализа. Иерархическая кластеризация. Числовые характеристики кластерного разбиения.
•	4 Задачи классификации	Байесовский подход к задаче классификации. Многомерное гауссовское распределение. Линейный и квадратичный дискриминантный анализ.
•	5 Регрессионный анализ	Однофакторная линейная регрессионный. Множественная линейная регрессия. Проверка гипотез о коэффициентах регрессии. Нелинейная регрессия. Логистическая регрессия.

## 6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Рекомендуемая тематика учебных занятий в форме контактной работы

Рекомендуемая тематика учебных занятий лекционного типа (предусматривающих преимущественную передачу учебной информации преподавателями):

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
-------	--	-------------------------------------

•	1 Основные понятия интеллектуального анализа данных	Классификация задач интеллектуального анализа данных. Расстояния, сходства, метрики для различных типов данных. Этапы подготовки данных. Выдвижение гипотез. Методы сбора и систематизации фактов. Методы проведения экспертиз для выявления наиболее значимых факторов. Понятия парциальной и комплексной обработки. Анализ качества полученных моделей.
•	2 Использование NumPy и Pandas для анализа данных	Введение в NumPy. Векторизованные вычисления. Агрегирующие функции. Broadcasting. Логические операции. Введение в pandas. Индексы в pandas. Иерархические (множественные) индексы. Иерархические колонки. Агрегирование в pandas. Объединение нескольких наборов данных. Решение базовых задач анализа данных.
•	3 Кластерный анализ	Эвристические алгоритмы кластерного анализа. Иерархическая кластеризация. Числовые характеристики кластерного разбиения.
•	4 Задачи классификации	Байесовский подход к задаче классификации. Многомерное гауссовское распределение. Линейный и квадратичный дискриминантный анализ.
•	5 Регрессионный анализ	Однофакторная линейная регрессионный. Множественная линейная регрессия. Проверка гипотез о коэффициентах регрессии. Нелинейная регрессия. Логистическая регрессия.

### **Рекомендуемая тематика *практических* занятий:**

Тематика практических занятий совпадает с темами лекционных занятий.

На практических занятиях решаются задачи по теме занятия.

### **Требования к самостоятельной работе обучающихся**

1. Работа с лекционным материалом, предусматривающая проработку конспекта лекций и учебной литературы, по всем темам из п. 6 настоящей рабочей программы.
2. Выполнение домашнего задания, предусматривающего решение задач, выполнение упражнений, выдаваемых на практических занятиях, по всем темам из п. 6 настоящей рабочей программы.

Руководствуясь положениями статьи 47 и статьи 48 Федерального закона от 29 декабря 2012 г. N 273-ФЗ «Об образовании в Российской Федерации» научно-педагогические работники и иные лица, привлекаемые университетом к реализации данной образовательной программы, пользуются предоставленными академическими правами и свободами в части свободы преподавания, свободы от вмешательства в профессиональную деятельность; свободы выбора и использования педагогически обоснованных форм, средств, методов обучения и воспитания; права на творческую инициативу, разработку и применение авторских программ и методов обучения и воспитания в пределах реализуемой образовательной программы и отдельной дисциплины.

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане).

Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы.

## **7. Методические рекомендации по видам занятий**

Лекционные занятия.

В ходе лекционных занятий обучающимся рекомендуется выполнять следующие действия. Вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации по их применению. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Желательно оставить в рабочих конспектах поля, на которых во внеаудиторное время можно сделать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений.

Практические и семинарские занятия.

На практических и семинарских занятиях в зависимости от темы занятия выполняется поиск информации по решению проблем, практические упражнения, контрольные работы, выработка индивидуальных или групповых решений, итоговое

обсуждение с обменом знаниями, участие в круглых столах, разбор конкретных ситуаций, командная работа, представление портфолио и т.п.

Самостоятельная работа.

Самостоятельная работа осуществляется в виде изучения литературы, эмпирических данных по публикациям и конкретных ситуаций из практики, подготовке индивидуальных работ, работа с лекционным материалом, самостоятельное изучение отдельных тем дисциплины; поиск и обзор литературы и электронных источников; чтение и изучение учебника и учебных пособий.

## 8. Фонд оценочных средств

### 8.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы в рамках учебной дисциплины

Основными этапами формирования указанных компетенций при изучении обучающимися дисциплины являются последовательное изучение содержательно связанных между собой тем учебных занятий. Изучение каждой темы предполагает овладение обучающимися необходимыми компетенциями. Результат аттестации обучающихся на различных этапах формирования компетенций показывает уровень освоения компетенций.

Контролируемые разделы (темы) дисциплины	Индекс контролируемой компетенции (или её части)	Оценочные средства по этапам формирования компетенций
		текущий контроль по дисциплине
1 Основные понятия интеллектуального анализа данных	ПК-7	Опрос
2 Использование NumPy и Pandas для анализа данных	ПК-7	Опрос
3 Кластерный анализ	ПК-7	Опрос
4 Задачи классификации	ПК-7	Опрос
5 Регрессионный анализ	ПК-7	Опрос

### 8.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности в процессе текущего контроля

В качестве оценочных средств текущего контроля успеваемости предусмотрены:

решение индивидуальных заданий

Пример индивидуального задания

1. Содержательная задача

Набор данных содержит информацию о клиентах банка и предыдущей маркетинговой кампании банка.

Банк хочет повысить свой депозитный портфель путем предложения предыдущим клиентам срочного вклада.

На входе мы имеем профиль клиента, которому уже предлагался вклад и каков был результат предложения: сделал или не сделал вклад.

Цель заключается в том, чтобы предсказать, откликнется ли клиент на предложение банка по срочному вкладу или нет.

Желательно минимизировать количество предложений вклада, которые заведомо не будут приняты клиентом так как это повышает расходы на маркетинг кроме того большое количество неэффективных предложений может привести к репутационным потерям банка.

## 2. Данные

Этот набор данных взят из общедоступного источника.

[Moro et al., 2014] S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, In press, <http://dx.doi.org/10.1016/j.dss.2014.03.001>

Авторы: Sérgio Moro (ISCTE-IUL), Paulo Cortez (Univ. Minho) and Paulo Rita (ISCTE-IUL) @ 2014

Варианты данных: Банковский маркетинг (bank-full.csv)

Подробности и описание <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

## 3. Формальная задача

Дано: данные маркетингового исследования (профиль клиента) и результат взаимодействия клиента с банком.

Цель бинарной классификации - предсказать, подпишет ли клиент банковский депозит (переменная  $y$ ).

Показатель качества: минимизация неэффективных предложений банковского продукта.

Необходимо определить какой из классификаторов (LinearDiscriminantAnalysis, QuadraticDiscriminantAnalysis, KNeighborsClassifier, LogisticRegression, GaussianNB) дает наилучшую модель машинного обучения.

Содержание отчета (в формате PDF):

ФИО полностью, факультет, курс, группа

Постановка задачи

Описание классификаторов (кратко)

Текст программы

Пример работы программы в виде скриншота + графики и диаграммы

Выводы по результатам выбора классификатора

### 8.3. Перечень вопросов и заданий для промежуточной аттестации по дисциплине

Вопросы к зачету

- Классификация задач интеллектуального анализа данных.
- Расстояния, сходства, метрики для различных типов данных.
- Этапы подготовки данных. Выдвижение гипотез.
- Методы сбора и систематизации фактов.
- Методы проведения экспертиз для выявления наиболее значимых факторов.
- Понятия парциальной и комплексной обработки.
- Анализ качества полученных моделей.
- Введение в NumPy. Векторизованные вычисления. Агрегирующие функции. Broadcasting. Логические операции.
- Введение в pandas. Индексы в pandas. Иерархические (множественные) индексы. Иерархические колонки.
- Агрегирование в pandas. Объединение нескольких наборов данных. Решение базовых задач анализа данных.
- Эвристические алгоритмы кластерного анализа. Иерархическая кластеризация. Числовые характеристики кластерного разбиения.
- Байесовский подход к задаче классификации. Многомерное гауссовское распределение. Линейный и квадратичный дискриминантный анализ.
- Однофакторная линейная регрессионный.
- Множественная линейная регрессия.
- Проверка гипотез о коэффициентах регрессии.
- Нелинейная регрессия. Логистическая регрессия.

### 8.4. Планируемые уровни сформированности компетенций обучающихся и критерии оценивания

Уровни	Содержательное описание уровня	Основные признаки выделения уровня (этапы формирования компетенции, критерии оценки сформированности)	Пятибалльная шкала (академическая) оценка	Двухбалльная шкала, зачет	БРС, % освоения (рейтинговая оценка)
Повышенный	Творческая деятельность	<i>Включает нижестоящий уровень.</i>	отлично	зачтено	86-100

		Умение самостоятельно принимать решение, решать проблему/задачу теоретического и прикладного характера на основе изученных методов, приемов, технологий			
Базовый	Применение знаний и умений в более широких контекстах учебной и профессиональной деятельности, нежели по образцу с большей степени самостоятельности и инициативы	<i>Включает нижестоящий уровень.</i> Способность собирать, систематизировать, анализировать и грамотно использовать информацию из самостоятельно найденных теоретических источников и иллюстрировать ими теоретические положения или обосновывать практику применения	хорошо		71-85
Удовлетворительный (достаточный)	Репродуктивная деятельность	Изложение в пределах задач курса теоретически и практически контролируемого материала	удовлетворительно		55-70
Недостаточный	Отсутствие признаков удовлетворительного уровня		неудовлетворительно	не зачтено	Менее 55

## **9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.**

### **Основная литература**

- Лонца, А. Алгоритмы обучения с подкреплением на Python : практическое руководство / А. Лонца ; пер. с англ. А. А. Слинкина. - Москва : ДМК Пресс, 2020. - 286 с. - ISBN 978-5-97060-855-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1210701> (дата обращения: 18.11.2023). – Режим доступа: по подписке.
- Э.Вирсански, Э. Генетические алгоритмы на Python : практическое пособие / Э. Вирсански ; пер. с англ. А. А. Слинкина. - Москва : ДМК Пресс, 2020. - 286 с. - ISBN 978-5-97060-857-9. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1210703> (дата обращения: 18.11.2023). – Режим доступа: по подписке.

### Дополнительная литература

- Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 130 с. — ISBN 978-5-00101-908-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/151502> (дата обращения: 01.08.2022). — Режим доступа: для авториз. пользователей.
- Гаврилова, И. В. Основы искусственного интеллекта : учебное пособие / И. В. Гаврилова, О. Е. Масленникова. — 3-е изд., стер. — Москва : ФЛИНТА, 2019. — 283 с. — ISBN 978-5-9765-1602-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/115839> (дата обращения: 01.08.2022). — Режим доступа: для авториз. пользователей.
- Джонс, М. Т. Программирование искусственного интеллекта в приложениях : практическое пособие : [16+] / М. Т. Джонс ; пер. с англ. А. И. Осипова. — Москва : ДМК Пресс, 2018. — 313 с. : табл., схем. — Режим доступа: по подписке. (дата обращения: 01.08.2022). — ISBN 978-5-97060-579-0. — Текст : электронный.
- Жданов, А. А. Автономный искусственный интеллект : учебное пособие / А. А. Жданов. — 5-е изд. — Москва : Лаборатория знаний, 2020. — 362 с. — ISBN 978-5-00101-655-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/135544> (дата обращения: 01.08.2022). — Режим доступа: для авториз. пользователей.
- Остроух, А. В. Системы искусственного интеллекта : монография / А. В. Остроух, Н. Е. Суркова. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 228 с. — ISBN 978-5-8114-8519-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/176662> (дата обращения: 01.08.2022). — Режим доступа: для авториз. пользователей.

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

- НЭБ Национальная электронная библиотека, диссертации и прочие издания
- eLIBRARY.RU Научная электронная библиотека, книги, статьи, тезисы докладов конференций
- Гребенников Электронная библиотека ИД журналы
- ЭБС Консультант студента
- ПРОСПЕКТ ЭБС
- ЭБС ZNANIUM.COM
- ЭБС IBOOKS.RU
- РГБ Информационное обслуживание по МБА
- БЕН РАН
- Электронно-библиотечная система (ЭБС) Кантиана (<https://elib.kantiana.ru/>)

### 11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине.

Программное обеспечение обучения включает в себя:

- система электронного образовательного контента БФУ им. И. Канта – [www.lms.kantiana.ru](http://www.lms.kantiana.ru), обеспечивающую разработку и комплексное использование электронных образовательных ресурсов;
- серверное программное обеспечение, необходимое для функционирования сервера и связи с системой электронного обучения через Интернет;
- корпоративная платформа;
- установленное на рабочих местах обучающихся ПО: Microsoft Windows 7, Microsoft Office Standart 2010, антивирусное программное обеспечение Kaspersky Endpoint Security.
- СУБД PostgreSQL (Свободное ПО, лицензия - Freeware).
- MongoDB (Свободное ПО, лицензия - Freeware).
- Python 2.7.15 (Anaconda2 5.2.0 64-bit)
- Python 3.6.5 (Anaconda3 5.2.0 64-bit)

## **12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.**

Для проведения занятий лекционного типа, практических и семинарских занятий используются специальные помещения (учебные аудитории), оборудованные техническими средствами обучения – мультимедийной проекционной техникой. Для проведения занятий лекционного типа используются наборы демонстрационного оборудования.

Для проведения лабораторных работ, (практических занятий – при необходимости) используются специальные помещения (учебные аудитории), оснащенные специализированным лабораторным оборудованием: персональными компьютерами с возможностью выхода в интернет и с установленным программным обеспечением, заявленным в п.11.

Для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используются специальные помещения (учебные аудитории), оборудованные специализированной мебелью (для обучающихся), меловой / маркерной доской.

Для организации самостоятельной работы обучающимся предоставляются помещения, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.

Для обучения инвалидов и лиц с ограниченными возможностями здоровья университетом могут быть представлены специализированные средства обучения, в том числе технические средства коллективного и индивидуального пользования.

## **5. Программа практики**

«Не предусмотрена».

## **6. Программа итоговой аттестации по модулю**

Определение результатов освоения модуля на основе вычисления оценки по каждому элементу модуля.

Оценка по модулю рассчитывается по формуле:

$$R_j^{\text{мод}} = \frac{k_1 R_1 + k_2 R_2 + k_3 R_3 + \dots + k_n R_n + k_{\text{пр}} R_{\text{пр}} + R_{\text{кур}}}{k_1 + k_2 + k_3 + \dots + k_{\text{пр}}}$$

Где:

$R_j^{\text{мод}}$  – оценка по модулю

$k_1, k_2, k_3, \dots, k_n$  – зачетные единицы дисциплин, входящих в модуль

$k_{\text{пр}}$  – зачетные единицы по практике

$R_1, R_2, R_3, \dots, R_n$  – оценки по дисциплинам модуля

$R_{\text{пр}}$  – оценка по практике

$R_{\text{кур}}$  – оценка по курсовой работе

В случае, если по дисциплине предусмотрен зачет без оценки, то за оценку по дисциплине принимается «5».

В случае, если по модулю применяется балльно-рейтинговая система, то

$R_1, R_2, R_3, \dots, R_n$  – рейтинговые баллы студента по дисциплинам модуля

$R_{\text{пр}}$  – рейтинговые баллы студента по практике

$R_{\text{кур}}$  – рейтинговые баллы студента по курсовой работе