

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»



УТВЕРЖДАЮ

Директор по образовательной деятельности

*С.Т. Князев*

С.Т. Князев

«12» Октября 2021 г.

## РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля  
М.1.5

Модуль  
Машинное обучение и искусственный интеллект

Екатеринбург, 2021

<b>Перечень сведений о рабочей программе модуля</b>	<b>Учетные данные</b>
<b>Образовательная программа</b> Прикладной искусственный интеллект (Practical Artificial Intelligence)	<b>Код ОП</b> 09.04.02
<b>Направление подготовки</b> Информационные системы и технологии	<b>Код направления и уровня подготовки</b> 09.04.02

Области образования, в рамках которых реализуется модуль образовательной программы по СУОС УрФУ:

<b>№ п/п</b>	<b>Перечень областей образования, для которых разработан СУОС УрФУ</b>	<b>Уровень подготовки</b>
1.	Инженерное дело, технологии и технические науки	магистратура

Программа модуля составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Долганов Антон Юрьевич	кандидат технических наук, без ученого звания	Доцент	Кафедра радиоэлектроники и телекоммуникаций, ИРИТ-РТФ, УрФУ
2	Созыкин Андрей Владимирович	кандидат технических наук, нет	доцент	Кафедра информационных технологий и систем управления, ИРИТ-РТФ, УрФУ

**Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий-РТФ**

Протокол № 7 от 11.10.2021 г.

# 1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Машинное обучение и искусственный интеллект

## 1.1. Аннотация содержания модуля

Модуль состоит из дисциплин: «Машинное обучение» и «Спортивный анализ данных»  
Целью освоения дисциплины «Машинное обучение» является освоение студентами основных вопросов теории вероятности, методов оптимизации и стохастических процессов для дальнейшего применения в разработке алгоритмов машинного обучения.

Цель освоения дисциплины «Спортивный анализ данных» - ознакомить студентов с современной платформой для проведения соревнований "Kaggle .com". Студенты узнают, как использовать различные алгоритмы и методы анализа данных для решения конкретных прикладных задач. На практике рассматриваются все типы задач анализа данных: анализ табличных данных, анализ временных рядов, обработка естественного языка, обработка изображений. В курсе рассматриваются методы анализа базовых обученных моделей, выбора и проверки новых возможностей, оптимальные методы поиска лучшего алгоритма решения задачи.

## 1.2. Структура и объем модуля

Таблица 1

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах и часах
1.	Машинное обучение	6 з.е., 216 час.
2.	Спортивный анализ данных	3 з.е., 108 час.
ИТОГО по модулю:		9 з.е., 324 час.

## 1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	<i>отсутствуют</i>
Постреквизиты и корреквизиты модуля	<i>отсутствуют</i>

## 1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2.1

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения
1	2	3
Машинное обучение	ОПК 1. Способен формулировать и решать научно-исследовательские, технические, организационно-экономические и	ОПК-1. 3-1. Соотносить проблемную область с соответствующей областью фундаментальных и инженерных наук ОПК-1. 3-2. Привести примеры терминологии, принципов, методологических подходов и законов фундаментальных и инженерных наук,

	<p>комплексные задачи, применяя фундаментальные знания.</p>	<p>применимых для формулирования и решения задач проблемной области знания.</p> <p>ОПК-1. У-1. Использовать для формулирования и решения задач проблемной области терминологию, основные принципы, методологические подходы и законы фундаментальных и общеинженерных наук.</p> <p>ОПК-1. У-2. Критически оценить возможные способы решения задач проблемной области, используя знания фундаментальных и общеинженерных наук.</p> <p>ОПК-1. П-1. Работая в команде, разрабатывать варианты формулирования и решения научно-исследовательских, технических, организационно-экономических и комплексных задач, применяя знания фундаментальных и общеинженерных наук.</p> <p>ОПК-1. Д-1. Проявлять лидерские качества и умения командной работы.</p>
--	---	--

Таблица 2.2

Перечень дисциплин модуля	Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения
1	2	3	4
Машинное обучение	ОПК-8. Способен разрабатывать алгоритмы и программные средства для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. З-1. Знает инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-8.1. У-1. Умеет применять инструментальные среды, программно-технические платформы для решения профессиональных задач
		ОПК-8.2. Разрабатывает оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	ОПК-8.2. З-1. Знает принципы разработки оригинальных программных средств для решения профессиональных задач ОПК-8.2. У-1. Умеет разрабатывать оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта
	ПК-3. Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач	ПК-3.1. Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	ПК-3.1. З-1. Знает классы методов и алгоритмов машинного обучения ПК-3.1. У-1. Умеет ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения

		ПК-3.2. Руководит исследовательской группой по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	ПК 3.2. 3-1. Знает методы и критерии оценки качества моделей машинного обучения ПК 3.2. У-1. Умеет определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области
Спортивный анализ данных	ОПК-8. Способен разрабатывать алгоритмы и программные средства для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. 3-1. Знает инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-8.1. У-1. Умеет применять инструментальные среды, программно-технические платформы для решения профессиональных задач
	ПК-7. Способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых субтехнологий искусственного интеллекта в прикладных областях	ПК-7.3. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию перспективных направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)	ПК-7.3. 3-1. Знает современное состояние и перспективы развития новых направлений, методов и технологий в области искусственного интеллекта ПК-7.3. У-1. Умеет проводить анализ новых направлений, методов и технологий в области искусственного интеллекта и определять наиболее перспективные для различных областей применения

### 1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной форме.

## 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН МОДУЛЯ

**ПРОГРАММА МОДУЛЯ**  
**Машинное обучение и искусственный интеллект**

**РАЗДЕЛ 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН**  
**МОДУЛЯ**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ 1**  
**Машинное обучение**

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Долганов Антон Юрьевич	кандидат технических наук, без ученого звания	Доцент	Кафедра радиоэлектроники и телекоммуникаций, ИРИТ-РТФ, УрФУ

**Рекомендовано учебно-методическим советом института Институт радиоэлектроники и информационных технологий-РТФ**

Протокол № 7 от 11.10.2021 г.

## 2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ 1 Машинное обучение

### 2.1. Технологии обучения, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология;
- С применением электронного обучения на основе электронных учебных курсов, размещенных на LMS-платформах УрФУ.

### 2.2. Содержание дисциплины 1

Таблица 1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
МО_1.1	История машинного обучения и базовые понятия	Определение машинного обучения (МО). Развитие МО: основные исторические этапы. Классификация задач в МО. Базовые понятия в МО.
МО_1.2	Данные	Типы данных. Представление данных. Предварительная обработка. Базы данных
МО_1.3	Методы уменьшения размерности	Матрица ковариации. Метод Главных Компонент. Сингулярное разложение Матрицы.
МО_1.4	Кластеризация	Метрики расстояния. Кластеризация k-Средних (k-Means). Иерархическая кластеризация. Кластеризация DBSCAN
МО_1.5	Регрессия	Линейная Регрессия. Метод наименьших квадратов. Градиентный спуск. Регуляризация. Метрики моделей регрессии.
МО_1.6	Классификация	Типы задач классификации. Логистическая регрессия. Метрики классификации. Матрица ошибок
МО_1.7	Ближайшие соседи.	Классификатор k-ближайших соседей (k-nearest neighbors). Регрессия k-ближайших соседей. Neighborhood Component Analysis. Визуализация данных методом t-SNE

МО_1.8	Байесовские методы	Теорема Байеса. Наивный Байесовский классификатор. Дискриминантный Анализ. Линейный дискриминант Фишера
МО_1.9	Метод опорных векторов	Опорные вектора. Зазор (margin). Ядра. Kernel Trick. Применение метода опорных векторов в задачах классификации и регрессии.
МО_1.10	Деревья Решений	Применение деревьев решений для решения задач классификации и регрессии. Основные элементы деревьев решений.
МО_1.11	Ансамблевые методы	Методы усреднения. Бэггинг. Случайный Лес (Random Forest). Методы Бустинга. AdaBoost. Градиентный бустинг
МО_1.12	Применение методов машинного обучения	Получение Данных. Предварительная Обработка. Отбор значимых параметров (feature selection). Выбор Модели. Оценка Модели. Настройка модели (fine-tuning). Анализ Модели
МО_2.1	Основы Компьютерного Зрения	Цели и задачи компьютерного зрения. Цвет и его восприятие. Цветовые модели. Регистрация изображений. Библиотека OpenCV/
МО_2.2	Компьютерное Зрение: Базовые Операции	Коррекция артефактов изображений. Фильтры и Свертка. Преобразование Фурье.
МО_2.3	Выделение Объектов	Бинаризация. Поиск границ. Distance Transform. Сегментация. Поиск по шаблону
МО_2.4	Особые Точки на Изображениях	Определение особых точек. Детекторы и Детекторы. Детектор Харриса. Scale-Invariant Feature Transform. Speeded-Up Robust Features. Features from Accelerated Segment Test. Binary Robust Independent Elementary Features. Oriented FAST and Rotated BRIEF
МО_2.5	Применение Машинного обучения в Компьютерном Зрении	Задачи Машинного обучения в Компьютерном зрении. Мешок Слов. Визуальный словарь. Viola–Jones object detection framework.

МО_2.6	Нейронные Сети	Введение в нейронные сети. Описание TensorFlow. Полносвязные (Dense) нейронные сети. Функции потерь. Оптимизаторы. Регуляризация Нейронных сетей. Сверточные (Convolutional) нейронные сети. Pooling
МО_2.7	Архитектуры Нейронных сетей для классификации изображений	LeNet-5. AlexNet. VGG. GoogLeNet. ResNet. MobileNet. EfficientNet. Перенос обучения (Transfer Learning)
МО_2.8	Архитектуры Нейронных Сетей для Поиска Объектов	Классификация с локализацией. Region Based Convolutional Neural Networks. Архитектуры семейства YOLO. Single Shot MultiBox Detector. CenterNet.
МО_2.9	Генеративные модели Нейронных Сетей	Генеративно-состязательные сети. Энкодеры. Декодеры. Вариационные Авто-Энкодеры. Перенос Стиля.
МО_2.10	Введение в обработку естественного языка	Цели и задачи обработки естественного языка. Базовые понятия. Предварительная обработка текста. Sentiment Analysis. Векторные модели.
МО_2.11	Вероятностные Модели в обработке естественного языка	Автокоррекция слов. Марковские цепи. Автозаполнение. N-граммы. Языковые Модели.
МО_2.12	Применение нейронных сетей для обработки естественного языка	Последовательные модели. Embedding. Recurrent Neural Network. Gated Recurrent Units. Long Short-Term Memory. Внимание (Attention). Трансформеры

**2.3.** Программа дисциплины реализуется полностью на иностранном языке.

## **2.4. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ** Машинное обучение

### **Электронные ресурсы (издания)**

1. Курс Methods of Machine Learning <https://elearn.urfu.ru/enrol/index.php?id=5960> (дата обращения: 05.10.2021).

### Профессиональные базы данных, информационно-справочные системы

1. Цифровая библиотека научно-технических изданий Института инженеров по электротехнике и радиоэлектронике (Institute of Electrical and Electronic Engineers (IEEE)) на английском языке – <http://www.ieee.org/ieeexplore>
2. Oxford University Press – <http://www.oxfordjournals.org/en/>
3. Архив препринтов с открытым доступом – <https://arxiv.org/>

### Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

### Базы данных, информационно-справочные и поисковые системы

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
2. eBook Collections Springer Nature – <https://link.springer.com/>
3. Гугл Академия – <https://scholar.google.ru/>
4. Электронный научный архив УрФУ – <https://elar.urfu.ru/>
5. Зональная научная библиотека (УрФУ) – <http://lib2.urfu.ru/>
6. Портал информационно-образовательных ресурсов УрФУ – [study.urfu.ru](http://study.urfu.ru)
7. Электронно-библиотечная система «Лань» – [e.lanbook.com](http://e.lanbook.com)
8. Университетская библиотека ONLINE – [biblioclub.ru](http://biblioclub.ru)
9. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks) – [bibliocomplectator.ru/available](http://bibliocomplectator.ru/available)
10. Электронные информационные ресурсы Российской государственной библиотеки – [www.rsl.ru](http://www.rsl.ru)
11. Научная электронная библиотека – <http://elibrary.ru/>
12. Научная электронная библиотека «КиберЛенинка» – <https://cyberleninka.ru/>
13. Web of Science Core Collection – <http://apps.webofknowledge.com/>

## 2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Машинное обучение

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 2

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции, практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Браузер (Google Chrome, Mozilla Firefox)

		Персональные компьютеры по количеству обучающихся Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	
--	--	--	--

**ПРОГРАММА МОДУЛЯ**  
Машинное обучение и искусственный интеллект

**РАЗДЕЛ 2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН  
МОДУЛЯ**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ 2**  
Спортивный анализ данных

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Созыкин Андрей Владимирович	кандидат технических наук, нет	доцент	Кафедра информационных технологий и систем управления, ИРИТ-РТФ, УрФУ

**Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий-РТФ**

Протокол № 7 от 11.10.2021 г.

## 2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ 2 Спортивный анализ данных

### 2.1. Технологии обучения, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология;
- Смешанная модель обучения с использованием онлайн-курса УрФУ;
- Исключительно электронного обучения с использованием внутреннего онлайн-курса УрФУ.

### 2.2. Содержание дисциплины

Таблица 1

Код раздела, темы	Раздел, тема дисциплины	Содержание
1	Обзор инструментов применяемых для анализа данных и машинного обучения	Знакомство с инструментами: jupyter notebook, google colab, kaggle kernel. Базовый функционал и приемы работы. Основные понятия Data Science и Machine Learning
2	Обзор библиотеки Numpy	Применение библиотеки numpy для операция над тензорами и решения задач линейной алгебры и генерации тензоров различной размерности по заданным правилам
3	Работа с библиотекой Pandas (Matplotlib, seaborn)	Библиотека для работы с табличными данными pandas: чтение и запись файлов, методы просмотра данных, индексация, срезы, фильтрация, агрегация и сортировка данных, визуализация данных, разведывательный анализ данных, базовые статистики.
4	Модели машинного обучения для решения задач классификации	Постановка задачи классификации. Изучение алгоритмов классификации в машинном обучении: логистическая регрессия, машины опорных векторов, наивный Байес, K ближайших соседей, деревья решений, ансамблевые модели
5	Модели машинного обучения для решения задач регрессии	Постановка задачи регрессии. Изучение алгоритмов регрессии в машинном обучении: линейная регрессия, машины опорных векторов, K ближайших соседей, деревья решений, ансамблевые модели. Регуляризация моделей машинного обучения
6	Метрики качества при решении задач классификации и регрессии	Изучение метрик качества моделей машинного обучения для классификации и регрессии. Рассмотрение сильных и слабых сторон различных метрик для различных задач. Применение изученных метрик
7	Предварительная обработка данных при решении задач анализа данных	Методы предварительной обработки: заполнение пропусков в данных, обработка выбросов в данных.
8	Преобразование и создание новых признаков для решения задач анализа данных	Изменение непрерывных типов данных: логарифмирование, построение полиномов, дискретизация данных, стандартизация и нормализация данных. Различные методы преобразование категориальных данных.
9	Методы выбора лучших	Методы отбора лучших для решения поставленной

	признаков для решения задач классификации и регрессии	задачи признаков в задачах классификации и регрессии: на основе статистических подходов и методов машинного обучения
10	Методы выбора лучших моделей и алгоритмов для решения поставленных задач анализа данных	Приемы отбора лучших моделей для решения поставленных задач. Различные методики разбиения выборок для более стабильной и правдоподобной оценки
11	Методы подбора гипер параметров моделей машинного обучения и построения конвейеров автоматизации решения задач анализа данных	Инструменты построения конвейеров обработки данных в машинном обучении. Методы и инструменты автоматического подбора гипер параметров моделей машинного обучения, а также подбора самих моделей и методов предварительной обработки
12	Работа с временными рядами	Понятие временного ряда и его составляющие элементы. Методы прогнозирования на основе эконометрических подходов. Методы прогнозирования на основе регрессионного подхода и моделей машинного обучения. Генерация новых временных признаков для временного ряда.
13	Задачи обучения без учителя: понижение размерности	Обзор задач обучения без учителя и их применимость на практике. Методы понижения размерности и их практические приложения
14	Задачи обучения без учителя: кластеризация	Обзор методов кластеризации и их сравнительные характеристики. Методы оценки оптимального числа кластеров. Кластеризация для сегментации и анализа. Кластеризация, как метод сжатия информации и снижения размерности.
15	Задачи обучения без учителя: поиск аномалий	Постановка задачи поиска аномалий и выявления новизны. Методы поиска аномалий на основе подходов машинного обучения.
16	Обработка естественного языка: предварительная обработка	Знакомство с задачами обработки естественного языка.: Методы первичной предварительной обработки текстов: очистка, токенизация, лемматизация, стемминг.
17	Обработка естественного языка: векторные модели	Преобразование текста в векторное пространство. Статистические методы построения векторных моделей: мешок слов, ONE, tf-idf. Тематическое моделирование: латентное размещение Дирихле (LDA), латентный семантический анализ (LSA). Нейросетевые модели: word2vec, fastText, GloVe.
18	Обработка естественного языка: типы решаемых задач	Рассмотрение спектра задач обработки естественного языка: моделирование языка, классификация, поиск именованных сущностей, суммаризация, генерация.
19	Введение в нейронные сети	Базовые понятия нейронных сетей: нейрон, обучение нейрона, нейронная сеть, обучение нейронной сети. Функции потерь и метрики качества.
20	Нейронные сети для решения задач регрессии	Применение нейронных сетей для решения задач регрессии. Специфические для регрессии функции активации нейронов и метрики качества. Методы предварительной обработки данных.
21	Нейронные сети для решения задач классификации	Применение нейронных сетей для решения задач классификации. Специфические для классификации функции

		активации нейронов и метрики качества. Методы предварительной обработки данных.
22	Нейронные сети для решения задач обработки изображений	Знакомство с задачами обработки изображений. Изучение сверточных нейронных сетей. Понятие ядра, свертки и подвыборки.
23	Нейронные сети для решения задач обработки изображений: современные архитектуры	Изучение современных архитектур нейронных сетей для решения задач обработки изображений: VGG-16, VGG-19, ResNet, Inception, Xception, DenseNet, MobileNet, EfficientNet.
24	Нейронные сети для решения задач обработки естественного языка	Архитектуры нейронных сетей для решения задач обработки естественного языка: одномерные сверточные сети, рекуррентные нейронные сети.
25	Нейронные сети для решения задач обработки естественного языка: современные архитектуры	Обзор современных модификаций нейронных сетей для решения задач обработки естественного языка: трансформеры, механизм внимания, BERT
26	Нейронные сети для решения задач предсказания временных рядов	Применение нейронных сетей для решения задачи прогнозирования временных рядов: одномерные сверточные нейронные сети, рекуррентные нейронные сети. WaveNet для прогнозирования временного ряда

**2.3.** Программа дисциплины реализуется полностью на иностранном языке.

#### **2.4. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ Спортивный анализ данных**

##### **Электронные ресурсы (издания)**

1. Хливненко, Л. В. Практика нейросетевого моделирования : учебное пособие для вузов / Л. В. Хливненко, Ф. А. Пятакович. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 200 с. — ISBN 978-5-8114-8264-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/173811> (дата обращения: 05.10.2021).
2. Клетте, Р. Компьютерное зрение. Теория и алгоритмы : учебник / Р. Клетте ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 506 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131691> (дата обращения: 08.10.2021).
3. Тарасов, И. Е. Статистический анализ данных в информационных системах : учебно-методическое пособие / И. Е. Тарасов. — Москва : РТУ МИРЭА, 2020. — 96 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163854> (дата обращения: 06.10.2021).
4. Ростовцев, В. С. Искусственные нейронные сети : учебник / В. С. Ростовцев. — Санкт-Петербург : Лань, 2019. — 216 с. — ISBN 978-5-8114-3768-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/122180> (дата обращения: 08.10.2021). — Режим доступа: для авториз. пользователей.

## Профессиональные базы данных, информационно-справочные системы

1. Applied Science & Technology Source. EBSCO publishing <http://search.ebscohost.com>
2. Wiley Online Library <http://onlinelibrary.wiley.com/>
3. Гугл Академия <https://scholar.google.ru/>

### Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

### Базы данных, информационно-справочные и поисковые системы

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
2. eBook Collections Springer Nature – <https://link.springer.com/>
3. Гугл Академия – <https://scholar.google.ru/>
4. Электронный научный архив УрФУ – <https://elar.urfu.ru/>
5. Зональная научная библиотека (УрФУ) – <http://lib2.urfu.ru/>
6. Портал информационно-образовательных ресурсов УрФУ – [study.urfu.ru](http://study.urfu.ru)
7. Электронно-библиотечная система «Лань» – [e.lanbook.com](http://e.lanbook.com)
8. Университетская библиотека ONLINE – [biblioclub.ru](http://biblioclub.ru)
9. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks) – [bibliocomplectator.ru/available](http://bibliocomplectator.ru/available)
10. Электронные информационные ресурсы Российской государственной библиотеки – [www.rsl.ru](http://www.rsl.ru)
11. Научная электронная библиотека – <http://elibrary.ru/>
12. Научная электронная библиотека «КиберЛенинка» – <https://cyberleninka.ru/>
13. Web of Science Core Collection – <http://apps.webofknowledge.com/>

## 2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ 2

### Спортивный анализ данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 2

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции; Практические занятия.	Компьютерный класс. Мультимедийный проектор с экраном.; Сетевое оборудование.; Локальная сеть с выходом в глобальную сеть Internet.	MS PowerPoint, MS Word, MS Excel. MS Teams Используется бесплатно-распространяемое программное обеспечение: 1. Python – <a href="https://www.python.org/">https://www.python.org/</a>

## **ОЦЕНОЧНЫЕ МАТЕРИАЛЫ**

**Код модуля**  
М.1.5

**Модуль**  
Машинное обучение и искусственный интеллект

**Екатеринбург, 2021**

Оценочные материалы по модулю составлены авторами:

<b>№ п/п</b>	<b>Фамилия, имя, отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Долганов Антон Юрьевич	кандидат технических наук, без ученого звания	Доцент	Кафедра радиоэлектроники и телекоммуникаций, ИРИТ-РТФ, УрФУ

**Рекомендовано учебно-методическим советом института радиоэлектроники и информационных технологий - РтФ**

Протокол № 7 от 11 октября 2021 г.

## 1. СТРУКТУРА И ОБЪЕМ МОДУЛЯ МАШИННОЕ ОБУЧЕНИЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах и часах	Форма итоговой промежуточной аттестации по дисциплинам модуля и в целом по модулю
1.	Машинное обучение	6/216	Зачёт, экзамен
2.	Спортивный анализ данных	3/108	Зачет
<b>ИТОГО по модулю:</b>		<b>9/324</b>	

## 2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО МОДУЛЮ

Не предусмотрено

### Раздел 3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

Машинное обучение

Модуль М 1.5 Машинное обучение и искусственный интеллект

Оценочные материалы составлены автором(ами):

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Долганов Антон Юрьевич	кандидат технических наук, без ученого звания	Доцент	Кафедра радиоэлектроники и телекоммуникаций, ИРИТ-РТФ, УрФУ

# 1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (ИНДИКАТОРЫ) ПО ДИСЦИПЛИНЕ МОДУЛЯ МАШИННОЕ ОБУЧЕНИЕ

Таблица 1.1

Код и наименование компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3
ОПК 1. Способен формулировать и решать научно-исследовательские, технические, организационно-экономические и комплексные задачи, применяя фундаментальные знания.	<p>ОПК-1. 3-1. Соотносить проблемную область с соответствующей областью фундаментальных и инженерных наук</p> <p>ОПК-1. 3-2. Привести примеры терминологии, принципов, методологических подходов и законов фундаментальных и инженерных наук, применимых для формулирования и решения задач проблемной области знания.</p> <p>ОПК-1. У-1. Использовать для формулирования и решения задач проблемной области терминологию, основные принципы, методологические подходы и законы фундаментальных и инженерных наук.</p> <p>ОПК-1. У-2. Критически оценить возможные способы решения задач проблемной области, используя знания фундаментальных и инженерных наук.</p> <p>ОПК-1. П-1. Работая в команде, разрабатывать варианты формулирования и решения научно-исследовательских, технических, организационно-экономических и комплексных задач, применяя знания фундаментальных и инженерных наук.</p> <p>ОПК-1. Д-1. Проявлять лидерские качества и умения командной работы.</p>	Контрольная работа; Домашняя работа; Зачёт; Экзамен

Таблица 1.2

Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3	4
ОПК-8. Способен разрабатывать алгоритмы и программные средства для решения задач в области создания и применения искусственного	ОПК-8.1. Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного	<p>ОПК-8.1. 3-1. Знает инструментальные среды, программно-технические платформы для решения профессиональных задач</p> <p>ОПК-8.1. У-1. Умеет применять инструментальные среды, программно-технические</p>	Контрольная работа; Домашняя работа; Зачёт; Экзамен

интеллекта	интеллекта	платформы для решения профессиональных задач	
	ОПК-8.2. Разрабатывает оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	ОПК-8.2. 3-1. Знает принципы разработки оригинальных программных средств для решения профессиональных задач ОПК-8.2. У-1. Умеет разрабатывать оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	Контрольная работа; Домашняя работа; Зачёт; Экзамен
ПК-3. Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач	ПК-3.1. Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	ПК-3.1. 3-1. Знает классы методов и алгоритмов машинного обучения ПК-3.1. У-1. Умеет ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения	Контрольная работа; Домашняя работа; Зачёт; Экзамен
	ПК-3.2. Руководит исследовательской группой по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	ПК 3.2. 3-1. Знает методы и критерии оценки качества моделей машинного обучения ПК 3.2. У-1. Умеет определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области	Контрольная работа; Домашняя работа; Зачёт; Экзамен

## 2. ВИДЫ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ, ВКЛЮЧАЯ МЕРОПРИЯТИЯ ТЕКУЩЕЙ АТТЕСТАЦИИ

### 2.1. Распределение объема времени по видам учебной работы

Таблица 2

№ п/ п	Наименование дисциплины модуля Приложения искусственного интеллекта	Объем времени, отведенный на освоение дисциплины модуля								
		Аудиторные занятия, час.				Промежуточная аттестация (форма итогового контроля)	Контактная работа (час.)	Самостоятельная работа студента, включая текущую аттестацию (час.)	Всего по дисциплине	
		Занятия лекцион ного типа	Практиче ские работы	Лаборатор ные работы	Всего				Час.	Зач. ед.
1	2	3	4	5	6	7	8	9	10	11
1.	Машинное обучение	54	54	0	108	Зачет, экзамен	126,78	89,22	216	6

## 2.2.Виды СРС, количество и объем времени на контрольно-оценочные мероприятия СРС по дисциплине

Контрольно-оценочные мероприятия СРС включают самостоятельное изучение материала, подготовку к аудиторным занятиям и мероприятиям текущего контроля, выполнение и оформление внеаудиторных мероприятий текущего контроля и подготовку к мероприятиям промежуточного контроля.

Таблица 3

№ п/п	Вид самостоятельной работы студента по дисциплине модуля	Количество контрольно-оценочных мероприятий СРС	Объем контрольно-оценочных мероприятий СРС (час.)
1.	Подготовка к лекционным, практическим занятиям		27 час.
2.	Выполнение и оформление мероприятий текущего контроля:		
2.1	Контрольная работа	2	10 час.
2.2	Домашняя работа	5	25 час.
3.	Подготовка к зачету	1	12 час.
4.	Подготовка к экзамену	1	12 час.
5.	Самостоятельное изучение материала для подготовки к выполнению контрольных мероприятий		3,22 час.
Итого на СРС по дисциплине:			89,22 час.

## 2. ПРОЦЕДУРЫ КОНТРОЛЯ И ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ В РАМКАХ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ МОДУЛЯ В БАЛЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЕ (ТЕХНОЛОГИЧЕСКАЯ КАРТА БРС)

### 3.1. Процедуры текущей и промежуточной аттестации по дисциплине

<b>1 семестр</b>		
<b>1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.4</b>		
<b>Текущая аттестация на лекциях</b>	<b>Сроки – семестр, учебная неделя</b>	<b>Максимальная оценка в баллах</b>
Контрольная работа 1	I, 15	100
<b>Весовой коэффициент значимости результатов текущей аттестации по лекциям – 1.0</b>		
<b>Промежуточная аттестация по лекциям – нет</b>		
<b>Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.0</b>		
<b>2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0.6</b>		
<b>Текущая аттестация на практических/семинарских занятиях</b>	<b>Сроки – семестр, учебная неделя</b>	<b>Максимальная оценка в баллах</b>
Домашняя работа 1	I, 5	30
Домашняя работа 2	I, 10	30
Домашняя работа 3	I, 15	40
<b>Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям – 0.5</b>		
<b>Промежуточная аттестация по практическим/семинарским занятиям – зачет</b>		
<b>Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям – 0.5</b>		
<b>2 семестр</b>		
<b>1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.4</b>		
<b>Текущая аттестация на лекциях</b>	<b>Сроки – семестр, учебная неделя</b>	<b>Максимальная оценка в баллах</b>
Контрольная работа 2	II, 15	100

<b>Весовой коэффициент значимости результатов текущей аттестации по лекциям – 1.0</b>		
<b>Промежуточная аттестация по лекциям – нет</b> <b>Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.0</b>		
<b>2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0.6</b>		
<b>Текущая аттестация на практических/семинарских занятиях</b>	<b>Сроки – семестр, учебная неделя</b>	<b>Максимальная оценка в баллах</b>
Домашняя работа 4	II, 10	50
Домашняя работа 5	II, 15	50
<b>Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– 0.5</b>		
<b>Промежуточная аттестация по практическим/семинарским занятиям–экзамен</b> <b>Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям– 0.5</b>		

### 3. КРИТЕРИИ И УРОВНИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ МОДУЛЯ

4.1. В рамках БРС применяются утвержденные на кафедре/институте критерии (признаки) оценивания достижений студентов по дисциплине модуля (табл. 4) в рамках контрольно-оценочных мероприятий на соответствие указанным в табл.1 результатам обучения (индикаторам).

Таблица 4

#### Критерии оценивания учебных достижений обучающихся

<b>Результаты обучения</b>	<b>Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам</b>
Знания	Студент демонстрирует знания и понимание в области изучения на уровне указанных индикаторов и необходимые для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Умения	Студент может применять свои знания и понимание в контекстах, представленных в оценочных заданиях, демонстрирует освоение умений на уровне указанных индикаторов и необходимых для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.

Опыт /владение	Студент демонстрирует опыт в области изучения на уровне указанных индикаторов.
Личностные качества	Студент демонстрирует ответственность в освоении результатов обучения на уровне запланированных индикаторов. Студент способен выносить суждения, делать оценки и формулировать выводы в области изучения. Студент может сообщать преподавателю и коллегам своего уровня собственное понимание и умения в области изучения.

4.2. Для оценивания уровня выполнения критериев (уровня достижений обучающихся при проведении контрольно-оценочных мероприятий по дисциплине модуля) используется универсальная шкала (табл. 5).

Таблица 5

#### Шкала оценивания достижения результатов обучения (индикаторов) по уровням

Характеристика уровней достижения результатов обучения (индикаторов)				
№ п/п	Содержание уровня выполнения критерия оценивания результатов обучения (выполненное оценочное задание)	Шкала оценивания		
		Традиционная характеристика уровня		Качественная характеристика уровня
1.	Результаты обучения (индикаторы) достигнуты в полном объеме, замечаний нет	Отлично (80-100 баллов)	Зачтено	Высокий (В)
2.	Результаты обучения (индикаторы) в целом достигнуты, имеются замечания, которые не требуют обязательного устранения	Хорошо (60-79 баллов)		Средний (С)
3.	Результаты обучения (индикаторы) достигнуты не в полной мере, есть замечания	Удовлетворительно (40-59 баллов)		Пороговый (П)
4.	Освоение результатов обучения не соответствует индикаторам, имеются существенные ошибки и замечания, требуется доработка	Неудовлетворительно (менее 40 баллов)	Не зачтено	Недостаточный (Н)
5.	Результат обучения не достигнут, задание не выполнено	Недостаточно свидетельств для оценивания		Нет результата

#### 4. СОДЕРЖАНИЕ КОНТРОЛЬНО-ОЦЕНОЧНЫХ МЕРОПРИЯТИЙ ПО ДИСЦИПЛИНЕ МОДУЛЯ

Задания по контрольно-оценочным мероприятиям в рамках текущей и промежуточной аттестации должны обеспечивать освоение и достижение результатов обучения (индикаторов) и предметного содержания дисциплины на соответствующем уровне.

## 5.1. Описание контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

### 5.1.1. Практические занятия

Номер занятия	Примерный перечень тем практических работ
1.1	Vector and Matrix Operations Using the NumPy Library
1.2	Working with Datasets using the Pandas library
1.3	Decomposition of matrices using principal component method and singular value decomposition matrix
1.4	Clustering: k-Means, DBSCAN, hierarchical clustering
1.5	Implementation of the linear regression algorithm
1.6	Implementation of the logistic regression algorithm
1.7	Support Vector Machine Classification and Regression
1.8	K-Nearest Neighbors Classification and Regression
1.9	Data visualization using Neighborhood Component Analysis
1.10	T-SNE data visualization
1.11	Data classification using the Naive Bayesian classifier
1.12	Data classification using Discriminant Analysis
1.13	Data visualization using Fisher's Linear Discriminant
1.14	Data Classification and Regression Using Decision Trees
1.15	Data classification and regression using ensemble methods

2.1	Image preprocessing and artifact removal
2.2	Object detection: binarization, segmentation, template search
2.3	Detectors and Descriptors: SIFT, SURF, ORB
2.4	Bag of Words methods for Image Classification
2.5	Viola–Jones object detection
2.6	Dense Neural Networks in TensorFlow
2.7	Convolutional Neural Networks in TensorFlow
2.8	Transfer Learning in TensorFlow
2.9	Object Detection by Neural Networks in TensorFlow
2.10	Style Transfer in TensorFlow
2.11	Vector Models in Natural Language Processing
2.12	Autocorrection models
2.13	Text generation using Recurrent Neural Networks
2.14	Transfer Learning of Transformers

### **5.1.2. Лабораторные занятия**

Не предусмотрено

### **5.1.3. Курсовая работа / Курсовой проект**

*Не предусмотрено*

#### 5.1.4. Контрольная работа

Примерная тематика контрольных работ:

Номер работы	Примерная тематика контрольных работ
1.1	History of machine learning and basic concepts
1.2	Data
1.3	Matrix decomposition methods
1.4	Clustering
1.5	Regression
1.6	Classification
1.7	Nearest Neighbors
1.8	Bayesian methods
1.9	support vectors machines
1.10	Decision Trees
1.11	Ensemble methods
1.12	Best practices of machine learning application
2.1	Foundations of Computer Vision
2.2	Basic operations of Computer Vision
2.3	Object Detection

<b>2.4</b>	Key points on the Images
<b>2.5</b>	Application of Machine Learning in Computer Vision Tasks
<b>2.6</b>	Dense Neural Networks
<b>2.7</b>	Convolutional Neural Networks
<b>2.8</b>	Architectures of Neural Networks for Image Classification
<b>2.9</b>	Architectures of Neural Networks for Object Detection
<b>2.10</b>	Generative Neural Networks Models
<b>2.11</b>	Introduction to Natural Language Processing
<b>2.12</b>	Probability models in Natural Language Processing
<b>2.13</b>	Application of Neural Networks in Natural Language Processing

**Примерные задания** в составе контрольных работ:

- 1.1.1 Describe the difference between the machine learning approach and traditional programming
- 1.1.2 Describe the difference between supervised learning and unsupervised learning
- 1.2.1 Give couple examples of continuous and discrete data
- 1.2.2 You have three matrices A, B, C: A has dimensions  $5 \times 4$ , B has dimensions  $4 \times 6$ , C has dimensions  $3 \times 5$ . Please write all possible matrices which can be multiplied and write dimensions of the result(-s)
- 1.3.1 Explain, why it is necessary to perform data standardization before application of the Principal Components Analysis
- 1.3.2 In Principle Components Analysis what do eigenvalues and eigenvectors of the covariance matrix means?
- 1.3.3 In Singular Value Decomposition what do matrices U, S and V mean?
- 1.4.1 What is the relation between Euclidean distance and Minkowski distance?
- 1.4.2 What hyperparameter you need to set up for the k-means clustering algorithm?
- 1.4.3 Can silhouette coefficient be equal to negative number? If "yes" - in what cases, if "no" - why?
- 1.4.4 Hierarchical Clustering: what is the difference between different linkages?
- 1.4.5 DBSCAN: what points are considered to be Noise, Border Point, Core Point?

- 1.5.1 What are the main differences between Ordinary Least Squares and Gradient Descent for finding regression coefficients?
- 1.5.2 In what case Mean Square Logarithmic Error is more suitable metric than Mean Square Error?
- 1.5.3 Can Coefficient of Determination ( $R^2$  score) be equal to negative number? If "yes" - in what cases, if "no" - why?
- 1.5.4 Why L1 regularization can lead to reduction of features (in contrast to L2 regularization)?
- 1.6.1 What is the main difference between Classification tasks and Regression tasks?
- 1.6.2 Let's say the test results of the mr. C. came for coronavirus. The test gave a positive answer, although in fact mr. C. has no coronavirus. What kind of classification mistake did the test make?
- 1.6.3 Let's say there are two classifiers: the first classifier has Accuracy 95%, Sensitivity 99%, Specificity 50%; the second classifier has Accuracy 87%, Sensitivity 84%, Specificity 94%. What can you say about data used for classification? Which of these classifiers is more reliable?
- 1.6.4 How can one build the Decision surface for the Logistic Regression?
- 1.7.1 What is the main difference in using k-Nearest Neighbors for Classification and Regression
- 1.7.2 How can one find optimal value of hyperparameter k for Nearest Neighbors methods?
- 1.7.3 How can one make dimensionality reduction using Nearest Components Analysis?
- 1.7.4 What hyperparameter in t-SNE implementation is related to the balance between local and global aspects of data structure?
- 1.8.1 What kind of information you need to obtain in order to use Bayes theorem?
- 1.8.2 What does "Naive" in Naive Bayesian Classifier stand for?
- 1.8.3 Why Linear Discriminant analysis can be used as a dimensionality reduction technique?
- 1.8.4 What's the difference between Linear and Quadratic Discriminant Analysis?
- 1.9.1 What points are considered to be Support Vectors (for Classification and for Regression tasks)?
- 1.9.2 What is the difference between the Hard Margin SVM and Soft Margin SVM?
- 1.9.3 Why does Kernel trick help to improve results of the SVM?
- 1.10.1 What are the core elements of the Decision Tree?
- 1.10.2 How the most optimal decision nodes are selected?
- 1.10.3 What is the difference between using Decision Trees for Classification and Regression?
- 1.11.1 What is the difference between Boosting methods and Averaging methods?
- 1.11.2 What is the 'weak estimator' in context of Ensemble methods?
- 1.11.3 What is the difference in combining trees for Random Forest, Gradient Boosting and AdaBoost?
- 1.12.1 What is the core idea behind Filter Methods for feature selection?
- 1.12.2 What is the main principle behind Wrapper Methods for Feature selection?
- 1.12.3 What are the necessary Concepts which you need to successfully apply Genetic Algorithm for the Feature Selection?
- 2.1.1 What is the fundamental difference between RGB and CMYK color models?
- 2.1.2 What is the convenience of auxiliary models (HSI, HSV) versus RGB?
- 2.1.3 In which color model does the OpenCV library (cv2) open images by default?
- 2.2.1 Which of the filters can be presented as a convolution, and which cannot?
- 2.2.2 What is the maximum brightness value for a 4-bit image?
- 2.3.1 Which binarization method should you use to get the original object on a black background?
- 2.3.2 How can you get the total gradient of the image using vertical and horizontal filters?
- 2.3.3 Why does Canny's algorithm uses hysteresis (upper and lower thresholds)?
- 2.3.4 What is a Distance Transform essentially?
- 2.3.5 What idea is the Watershed segmentation algorithm based on?
- 2.4.1 What is the difference between detector and descriptor?

- 2.4.2 How can the second derivative of the Gaussian be used to predict scale of the blob?
- 2.4.3 What features are used as a descriptor in the method of SIFT?
- 2.4.4 What methods has the ORB method combined? Why can't they be fully used individually?
- 2.5.1 What computer vision problems can be solved using machine learning methods?
- 2.5.2 Describe the main steps in applying the Bag of Words method to image classification? What machine learning tasks are done on different steps?
- 2.5.3 What are the key points in the Viola Jones algorithm?
- 2.6.1 What is the main difference between the neural network approach and "classical" machine learning?
- 2.6.2 What are the key points of the backpropagation algorithm when training a neural network?
- 2.6.3 Why it is necessary to use activation functions in neural networks?
- 2.6.4 Why is it recommended to divide the initial sample into mini-batches when training a neural network?
- 2.6.5 What does the feedforward in context of neural networks mean?
- 2.7.1 What are the key methods used to avoid overfitting in the Neural Networks?
- 2.7.2 What is the idea behind RMSprop optimization?
- 2.7.3 What are trainable parameters when using Convolutional layers?
- 2.7.4 What are Pooling Layers used for?
- 2.7.5 What is the dimension of the neural network output if the original image was 30x30, it is fed to a convolutional layer with a 3x3 kernel, pooling with a 2x2 kernel and another 3x3 convolutional layer?
- 2.8.1 What does the 1x1 convolution do, and what are the weights that are learned during the training of the neural network?
- 2.8.2 What peculiarity of the ResNet architecture allowed to overcome Vanishing Gradient problem?
- 2.8.3 How is the amount of weights and operations is saved in the MobileNet architecture?
- 2.8.4 In which module of the tensorflow library are the ready-made pre-trained neural network architectures?
- 2.9.1 What is the peculiarity of the Image Classification with Localization?
- 2.9.2 What does the IOU (Intersection over Union) parameter show?
- 2.9.3 How does Non-maximal suppression in YOLO algorithm work?
- 2.9.4 Why YOLO 9000 was better, faster, stronger?
- 2.10.1 What is the difference between Generative and Discriminative Models of Machine Learning?
- 2.10.2 What is the basic principle of generative adversarial networks (GANs)?
- 2.10.3 What are the basic blocks of Auto-Encoders?
- 2.10.4 What is the peculiarity of Variational Auto-Encoders compared to basic Auto-Encoders?
- 2.11.1 What is the idea behind using Stemming as a Preprocessing technique?
- 2.11.2 What is the rationale behind the Naive machine translation using Vector models?
- 2.11.3 What are the core ideas of the Word2Vec vector space model?
- 2.12.1 How does the AutoCorrect algorithm balances between Candidate that needs less edits and Candidate which is more frequent in the Corpus?
- 2.12.2 What is the difference between the simple Markov Chains and Hidden Markov Chains in Natural Language processing tasks?
- 2.12.3 Why increasing N in N-gram models helps to improve quality of the generated text? How can you measure it?
- 2.13.1 What does the Embedding layer in Tensorflow library is actually doing?
- 2.13.2 What is the main idea behind the Recurrent Neural Networks?
- 2.13.3 What are the key improvements in Gated Recurrent Units?
- 2.13.4 What are key points in Long Short-Term Memory layers?
- 2.13.5 What are the Key, Values and Query in Attention layers?

## 2.13.6 Why Multi-Head attention improves results of the Transformers?

### 5.1.5. Домашняя работа

**Примерная тематика** домашних работ:

1. The basics of linear algebra in Python.
2. Basic machine learning algorithms.
3. Algorithms for clustering and data visualization in the sklearn library.
4. Regression algorithms in sklearn library.
5. Classification algorithms in sklearn library.
6. Advanced feature generation methods.
7. Applying Machine Learning Algorithms to Solve Dimension Reduction, Clustering, Regression and Classification problems.

**Примерные задания** в составе домашних работ:

Homework 1:

- a. perform text preprocessing
- b. perform sentiment analysis of the training corpus by means of classical machine learning methods
- c. apply probabilistic models for typos correction
- d. perform sentiment analysis of the training corpus by means of neural networks
- e. perform text generation by means of recurrent neural networks
- f. apply transfer learning of pre-trained Transformers Networks for Natural Language Processing

Homework 2:

- a. perform images classification using dense neural networks
- b. perform images classification using convolutional neural networks
- c. apply transfer learning of pre-trained convolutional neural networks for images classification
- d. apply transfer learning of pre-trained networks for object detection
- e. perform transfer of the style for images

Homework 3:

- a. perform image preprocessing by different methods
- b. perform object detection by means of basic computer vision tools
- c. perform key points detection
- d. apply machine learning for image classification

Homework 4:

- a. cluster the training dataset using the k-Means method
- b. cluster the training dataset using the DBSCAN method
- c. cluster the training dataset using the hierarchical clustering method
- d. visualize the training dataset using Neighborhood Component Analysis.
- e. visualize the training dataset using the t-SNE method

### 5.1.6. Расчетная работа

Не предусмотрено

### 5.1.7. Реферат / эссе / творческая работа

Не предусмотрено

### 5.1.8. Проектная работа

Не предусмотрено

### 5.1.9. Деловая (ролевая) игра / Дебаты / Дискуссия / Круглый стол

Не предусмотрено

### 5.1.10. Коллоквиум

Не предусмотрено

## 5.2. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля

### 5.2.1. Зачет в форме независимого тестового контроля (НТК)

НТК по дисциплине модуля не проводится.

**5.2.2. Зачет в традиционной форме** (устные /письменные ответы на вопросы экзаменационных билетов)

*Билет на зачет состоит из 5 вопросов, по одному на тематику.*

Тематика	Тема	Вопрос
1 Basic Concepts	1.1 Type of machine learning tasks	Provide all required definitions. Give examples of use cases
	1.2 Data Types	
	1.3 Underfitting and Overfitting	
	1.4 Gradient Descent	
	1.5 Cross-Validation	
	1.6 Confusion Matrix and Classification Metrics	
	1.7 Regression Metrics	
	1.8 Data Preprocessing	
2 Clustering	2.1 k-means clustering	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	2.2 Hierarchical clustering	
	2.3 DBSCAN	
3 Dimensionality Reduction	3.1 Principal Components Analysis	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	3.2 Singular Value Decomposition	
	3.3 Neighborhood Components Analysis	

	3.4 Data Visualization by t-SNE	
	3.5 Linear Discriminant Analysis	
4 Regression	4.1 Linear Regression	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	4.2 Regularization of linear regression	
	4.3 Regression by k-Nearest neighbors	
	4.4 Regression by Decision Trees	
	4.5 Regression by support vector machines	
	4.6 Regression by AdaBoost	
	4.7 Regression by Gradient Boosting	
	4.8 Regression by Random Forest	
5 Classification	5.1 Logistic Regression	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	5.2 Classification by k-Nearest neighbors	
	5.3 Naive Bayes classifier	
	5.4 Discriminant Analysis (Fisher's Interpretation)	
	5.5 Discriminant Analysis (Bayesian version)	
	5.6 Classification by Decision Trees	
	5.7 Classification by support vector machines (soft and hard margin)	
	5.8 Classification by support vector machines (kernel trick)	

	5.9 Classification by AdaBoost	
	5.10 Classification by Gradient Boosting	
	5.11 Classification by Random Forest	

#### Пример билета

1. Confusion Matrix and Classification Metrics. Provide all required definitions. Give examples of use cases
2. Hierarchical clustering. List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
3. Data Visualization by t-SNE. List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
4. Linear Regression. List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
5. Classification by AdaBoost. List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles

#### 5.2.3. Экзамен в традиционной форме (устные /письменные ответы на вопросы экзаменационных билетов)

Билет на экзамен состоит из 5 вопросов, по одному на тематику.

Тематика	Тема	Вопрос
<b>1 Core concepts</b>	1.1 Type of machine learning tasks	Provide all required definitions. Give examples of use cases
	1.2 Data Preprocessing	
	1.3 Underfitting and Overfitting	
	1.4 Gradient Descent	
	1.5 Cross-Validation	
	1.6 Confusion Matrix and Classification Metrics	
	1.7 Regression Metrics	
<b>2 Classic algorithms of machine learning</b>	2.1 PCA and SVD	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	2.2 k-means clustering	

	2.3 Affinity Propagation	
	2.4 Hierarchical clustering	
	2.5 DBSCAN	
	2.6 Linear Regression	
	2.7 Logistic regression	
	2.8 k-Nearest Neighbors	
	2.9 t-SNE	
	2.10 Neighborhood Component Analysis	
	2.11 Naive Bayes	
	2.12 Discriminant Analysis	
	2.13 Decision Trees	
	2.14 Support Vector Machines	
	2.15 AdaBoost	
	2.16 Gradient Boosting	
	2.17 Random Forest	
<b>3 Computer Vision</b>	3.1. Color models	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	3.2. Registration of images	
	3.3. Correction of image artifacts	
	3.4. Applying filters for image processing	
	3.5. Template matching	
	3.6. Edge detection in the image	
	3.7. Image segmentation	

	3.8. Feature points in the image and the Harris detector	
	3.9. SIFT detector and descriptor	
	3.10. SURF detector and descriptor	
	3.11. BRIEF detector and descriptor	
	3.12. Visual dictionary and Bag of Words Method	
	3.13. Viola-Jones algorithm for faces detection	
<b>4 Neural Networks</b>	4.1 Basic elements of dense neural networks: neurons, activation functions, weights and biases	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	4.2 Basic elements of convolutional neural networks: convolutions, padding, stride, pooling	
	4.3 Optimizers of neural networks: gradient descent, stochastic gradient descent, exponentially weighted averages, RMSProp, Adam	
	4.4 Peculiarities of neural network architectures for image classification	
	4.5 1x1 convolution and Google Inception neural network architecture for image classification	
	4.6 Residual blocks and ResNet neural network architecture for image classification	

	4.7 Depthwise Separable Convolution and MobileNet neural network architecture for image classification	
	4.8 Image classification with localization	
	4.9 Evolution of YOLO architectures for object detection	
	4.10 R-CNN Networks for object detection	
	4.11 Generative-Adversarial Networks	
	4.12 Auto-Encoders and Variational Auto-Encoders	
	4.13 Style transfer using neural networks	
<b>5 Natural Language Processing</b>	5.1 Preprocessing of the text in Python	List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
	5.2 Sentiment analysis by machine learning	
	5.3 Vector Space Model	
	5.4 Naïve Machine Translation	
	5.5 Autocorrect and edit distance	
	5.6 Part of speech tagging by Markov models	
	5.7 Autocomplete and N-Gram Language Models	
	5.8 Recurrent neural networks	
	5.9 Gated Recurrent Units and Long Short-Term Memory	

	5.10 Transformers: attention, multi-head attention, Encoders and Decoders	
	5.11 Transfer Learning in NLP and Language Models	

### **Пример билета**

- 1. Confusion Matrix and Classification Metrics.** Provide all required definitions. Give examples of use cases
- 2. Gradient Boosting.** List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
- 3. SIFT detector and descriptor.** List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
- 4. Residual blocks and ResNet neural network architecture for image classification.** List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles
- 5. Naïve Machine Translation.** List the main stages of the method application, indicate its strengths and weaknesses, opportunities and troubles

### **Раздел 3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ**

#### **СПОРТИВНЫЙ АНАЛИЗ ДАННЫХ**

##### **Модуль М.1.5 Машинное обучение и искусственный интеллект**

Оценочные материалы составлены автором(ами):

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Созыкин Андрей Владимирович	кандидат технических наук, нет	доцент	Кафедра информационных технологий и систем управления, ИРИТ-РТФ, УрФУ

## 1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (ИНДИКАТОРЫ) ПО ДИСЦИПЛИНЕ МОДУЛЯ Спортивный анализ данных

Таблица 1

Код и наименование компетенций, формируемые с участием дисциплины	Индикаторы достижения компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3	4
ОПК-8. Способен разрабатывать алгоритмы и программные средства для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного интеллекта	ОПК-8.1. 3-1. Знает инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-8.1. У-1. Умеет применять инструментальные среды, программно-технические платформы для решения профессиональных задач	Контрольная работа; домашняя работа; зачёт
ПК-7. Способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых субтехнологий искусственного интеллекта в прикладных областях	ПК-7.3. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию перспективных направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)	ПК-7.3. 3-1. Знает современное состояние и перспективы развития новых направлений, методов и технологий в области искусственного интеллекта ПК-7.3. У-1. Умеет проводить анализ новых направлений, методов и технологий в области искусственного интеллекта и определять наиболее перспективные для различных областей применения	Контрольная работа; домашняя работа; зачёт

## 2. ВИДЫ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ, ВКЛЮЧАЯ МЕРОПРИЯТИЯ ТЕКУЩЕЙ АТТЕСТАЦИИ

### 2.1. Распределение объема времени по видам учебной работы

Таблица 2

№ п/ п	Наименование дисциплины модуля	Объем времени, отведенный на освоение дисциплины модуля								
		Аудиторные занятия, час.				Промежуточная аттестация (форма итогового контроля)	Контактная работа (час.)	Самостоятельная работа студента, включая текущую аттестацию (час.)	Всего по дисциплине	
		Занятия лекцион ного типа	Практиче ские работы	Лаборатор ные работы	Всего				Час.	Зач. ед.
1	2	3	4	5	6	7	8	9	10	11
3.	Спортивный анализ данных	0	36	0	36	Зачет	41.65	66.35	108	3

### 3.1. Виды СРС, количество и объем времени на контрольно-оценочные мероприятия СРС по дисциплине

Контрольно-оценочные мероприятия СРС включают самостоятельное изучение материала, подготовку к аудиторным занятиям и мероприятиям текущего контроля, выполнение и оформление внеаудиторных мероприятий текущего контроля и подготовку к мероприятиям промежуточного контроля.

Таблица 3

№ п/п	Вид самостоятельной работы студента по дисциплине модуля	Количество контрольно-оценочных мероприятий СРС	Объем контрольно-оценочных мероприятий СРС (час.)
1.	Подготовка к аудиторным занятиям и мероприятиям текущего контроля: практическим занятиям.		13,5 час.
2.	Выполнение и оформление мероприятий текущего контроля:		
2.1	Контрольная работа	1	5 час.
2.2	Домашняя работа №1	1	5 час.
2.3	Домашняя работа №2	1	5 час.
2.4	Домашняя работа №3	1	5 час.
2.5	Домашняя работа №4	1	5 час.
8.	Подготовка к зачету	1	12 час.
9.	Самостоятельное изучение материала		15,85 час.
Итого на СРС по дисциплине:			66,35 час.

### 3. ПРОЦЕДУРЫ КОНТРОЛЯ И ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ В РАМКАХ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ МОДУЛЯ В БАЛЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЕ (ТЕХНОЛОГИЧЕСКАЯ КАРТА БРС)

#### 3.1. Процедуры текущей и промежуточной аттестации по дисциплине

<b>1. Лекции: не предусмотрены</b>
<b>2. Лабораторные занятия: не предусмотрены</b>

<b>3. Практические/семинарские занятия: коэффициент значимости совокупных результатов лабораторных занятий –1.0</b>			
<b>Текущая аттестация на практических/семинарских занятиях</b>	<b>Сроки – семестр, учебная неделя</b>	<b>Максимальная оценка в баллах</b>	
<i>Домашняя работа №1</i>	3 сем., 3 нед.	10	
<i>Домашняя работа №2</i>	3 сем., 6 нед.	10	
<i>Домашняя работа №3</i>	3 сем., 9 нед.	10	
<i>Домашняя работа №4</i>	3 сем., 12 нед.	10	
<i>Контрольная работа</i>	3 сем., 15 нед.	20	
<i>Выполнение и оформление практических работ</i>	3 сем.	20	
<i>Самостоятельное изучение материала</i>	3 сем.	20	
<b>Весовой коэффициент значимости результатов практическим/семинарским занятиям - 0,5</b>	<b>текущей</b>	<b>аттестации</b>	<b>по</b>
<b>Промежуточная аттестация по практическим/семинарским занятиям – зачет. Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям – 0,5</b>			

#### **4. КРИТЕРИИ И УРОВНИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ МОДУЛЯ**

4.2.В рамках БРС применяются утвержденные на кафедре/институте критерии (признаки) оценивания достижений студентов по дисциплине модуля (табл. 4) в рамках контрольно-оценочных мероприятий на соответствие указанным в табл.1 результатам обучения (индикаторам).

Таблица 4

##### **Критерии оценивания учебных достижений обучающихся**

<b>Результаты обучения</b>	<b>Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам</b>
Знания	Студент демонстрирует знания и понимание в области методов решения задач компьютерного зрения, в т.ч. понимание принципов получения и обработки цифровых изображений, классических методов решения задач компьютерного зрения, понимание основ глубокого обучения нейронных сетей и знания архитектур современных нейронных сетей в области компьютерного зрения.

Умения	Студент может применять свои знания и понимание в контекстах, представленных в заданиях представленных в рамках практических работ, а также в домашней работе. В т.ч. студент показывает умение использования изучаемых библиотек в рамках языка программирования Python, умение работы с цифровыми изображениями, умение работы с архитектурами глубокого обучения нейронных сетей.
Опыт /владение	Студент демонстрирует опыт в области изучения, достаточный для ответов на вопросы контрольных работы и вопросы по результатам заданий практических и домашней работы.
Личностные качества	Студент демонстрирует ответственность в освоении результатов обучения в т.ч. в выполнении контрольных работ, заданий на практиках и в выполнении домашней работы. Студент способен выносить суждения, делать оценки и формулировать выводы в области компьютерного зрения. Студент может сообщать преподавателю и коллегам своего уровня собственное понимание и умения в области компьютерного зрения.

4.3. Для оценивания уровня выполнения критериев (уровня достижений обучающихся при проведении контрольно-оценочных мероприятий по дисциплине модуля) используется универсальная шкала (табл. 5).

Таблица 5

#### Шкала оценивания достижения результатов обучения (индикаторов) по уровням

Характеристика уровней достижения результатов обучения (индикаторов)				
№ п/п	Содержание уровня выполнения критерия оценивания результатов обучения (выполненное оценочное задание)	Шкала оценивания		
		Традиционная характеристика уровня		Качественная характеристика уровня
1.	Результаты обучения (индикаторы) достигнуты в полном объеме, замечаний нет	Отлично (80-100 баллов)	Зачтено	Высокий (В)
2.	Результаты обучения (индикаторы) в целом достигнуты, имеются замечания, которые не требуют обязательного устранения	Хорошо (60-79 баллов)		Средний (С)
3.	Результаты обучения (индикаторы) достигнуты не в полной мере, есть замечания	Удовлетворительно (40-59 баллов)		Пороговый (П)
4.	Освоение результатов обучения не соответствует индикаторам,	Неудовлетворительно	Не зачтено	Недостаточный (Н)

	имеются существенные ошибки и замечания, требуется доработка	(менее 40 баллов)	
5.	Результат обучения не достигнут, задание не выполнено	Недостаточно свидетельств для оценивания	Нет результата

## 5. СОДЕРЖАНИЕ КОНТРОЛЬНО-ОЦЕНОЧНЫХ МЕРОПРИЯТИЙ ПО ДИСЦИПЛИНЕ МОДУЛЯ

Задания по контрольно-оценочным мероприятиям в рамках текущей и промежуточной аттестации должны обеспечивать освоение и достижение результатов обучения (индикаторов) и предметного содержания дисциплины на соответствующем уровне.

### 5.1. Описание контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

#### 5.1.1. Практические занятия

Номер занятия	Примерный перечень тем практических занятий
1	Overview of the tools used for data analysis and machine learning. Numpy library overview. Working with the Pandas library (matplotlib, seaborn)
2	Machine learning models for solving classification problems
3	Machine learning models for solving regression problems
4	Quality metrics in solving classification and regression problems
5	Data preprocessing when solving data analysis tasks. Transformation and creation of new features for solving data analysis problems
6	Methods for selecting the best features for solving classification and regression problems
7	Methods for selecting hyper parameters of machine learning models and building automation pipelines for solving data analysis problems
8	Working with time series
9	Tasks of teaching without a teacher: reducing the dimension Tasks of teaching without a teacher: finding anomalies
10	Natural Language Processing: preprocessing
11	Natural Language Processing: Vector models
12	Introduction to neural networks. Neural networks for solving regression problems
13	Neural networks for solving classification problems
14	Neural networks for solving image processing problems
15	Neural networks for solving image processing problems: modern architectures
16	Neural networks for solving natural language processing problems
17	Neural networks for solving natural language processing problems: modern architectures
18	Neural networks for solving time series prediction problems

### **5.1.2. Лабораторные занятия**

*Не предусмотрено*

### **5.1.3. Курсовая работа / Курсовой проект**

*Не предусмотрено*

### **5.1.4. Контрольная работа**

#### **Примерная тематика контрольных работ:**

*Контрольная работа №1*

Join the competition on the platform Kaggle.com, conduct a primary analysis of the competition data, perform preliminary data processing and, using various machine learning models, achieve the quality of the solution specified in the task condition. The problem of object classification is solved.

#### **Примерные задания в составе контрольных работ:**

Follow the link <https://www.kaggle.com/t/993ac81ad69a477580bd8cf1beca9a32>

There you will see a competition on the platform kaggle.com

On the "Overview" tab, you will see a description of the task.

The "Data" tab contains all the necessary data that will be used in the competition.

On the "Notebooks" tab in the "Public" section there is a basic solution with a sample code "baseline\_1\_5".

By copying the basic solution and running it at home, you will be able to achieve an accuracy of about 0.83.

Your task is to achieve the highest quality of the model by changing any parameters of the neural network architecture or neural network training parameters.

You can observe the result of your experiments on the "Leaderboard" tab immediately after uploading the response in the "My Submissions" tab.

Your goal is to achieve the quality of the model on test data 0.9 and higher. Then you pass the task

When using LeaderBoard, make sure that the result suits you, download the file of your solution, which you can see by opening your work notebook in viewing mode (but not editing) and going to the "Output" tab on the right.

Download the downloaded file in the ".csv" format here, on the course platform.

### **5.1.5. Домашняя работа**

*Домашняя работа №1:*

You are given a collaborative notebook with tasks for the numpy library. Complete all tasks and send a link to your solution to google forms

**Примерные тексты заданий:**

1. Create a one-dimensional numpy array of 10 elements filled with the number 1.5
2. Create a two dimensional numpy array of 25 elements filled with the number 0
3. Create a vector of length 15 with random integers ranging from -10 to 10. Reverse the vector
4. Create a 5 by 5 matrix with random integers ranging from -15 to 15. Replace the negative elements with their squares
5. Two matrices are given. Make sure they can be multiplied. Perform the matrix multiplication operation if possible, otherwise output an error

6. Given a vector, check if there is a Nan in it. Check if there is an infinity sign in it. Fill them with zeros

*Домашняя работа №2:*

You are given a coollaboratory notebook with tasks for the pandas library. Complete all tasks and send a link to your solution to google forms

***Примерные тексты заданий:***

A data set with demographic information about people receiving more than 50 thousand dollars a year or less is given. Analyze the set and answer the questions using the pandas library.

1. How many men and women (sex attribute) are represented in this dataset?
2. What is the average age (age attribute) of women?
3. What is the proportion of German citizens (a sign of native-country)? Is it true that people who receive more than 50k have at least a higher education? (attribute of education - Bachelors, Prof-school, Assoc-acdm, Assoc-voc, Masters or Doctorate)
5. Among whom is the share of earning a lot more (>50K): among married or single men (a sign of marital-status)? We consider married those whose marital-status begins with Married (Married-civ-spouse, Married-spouse-absent or Married-AF-spouse), the rest are considered single.
6. What is the maximum number of hours a person works per week (the hours-per-week attribute)? How many people work such a number of hours and what is the percentage of them earning a lot?

*Домашняя работа №3:*

You are given a coollaboratory notebook with tasks for the pandas, matplotlib, seaborn libraries. Complete all tasks and send a link to your solution to google forms

***Примерные тексты заданий:***

A data set with information about the Olympic Games for different years is given. Perform an initial visual analysis.

1. Plot graphs: linear, columnar, circular, histograms, boxes with whiskers - for each column of data,
2. Plot dot graphs reflecting the relationship between various features
3. Perform a correlation analysis and visualize it using a heat map
4. Make a brief description for each graph that reflects the information you saw on the visualization.

*Домашняя работа №4:*

You are given a coollaboratory notebook with tasks on sklearn libraries. Complete all tasks and send a link to your solution to google forms

***Примерные тексты заданий:***

A data set with information about the values of apartments, depending on their parameters, is given. Complete the following tasks.

1. Conduct a primary analysis: statistical and visual and present it.
2. Perform preliminary data processing: filling in gaps, processing outliers, generating new features.
3. Apply various machine learning models and achieve the quality of the model specified in the task according to the specified metric.

### **5.1.6. Расчетная работа / Расчетно-графическая работа**

*Не предусмотрено*

### **5.1.7. Реферат / эссе / творческая работа**

*Не предусмотрено*

### **5.1.8. Проектная работа**

*Не предусмотрено*

### **5.1.9. Деловая (ролевая) игра / Дебаты / Дискуссия / Круглый стол**

*Не предусмотрено*

### **5.1.10. Кейс-анализ**

*Не предусмотрено*

## **5.2. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля**

### **5.2.1. Зачет в форме независимого тестового контроля (НТК)**

НТК по дисциплине модуля не проводится.

### **5.2.2. Зачет в традиционной форме (устные / письменные ответы на вопросы):**

#### **Список примерных вопросов для зачета (3 семестр):**

1. Name the key data types that can be used in data mining tasks.
2. Name the tasks that can be solved when teaching machine learning models with a teacher. Give examples of such tasks.
3. Name the main stages of solving the problem of data mining.
4. Describe what kind of pre-processing and preparation of data needs to be carried out when working with tasks on images.
5. What tools and models are suitable for working with images?
6. Describe what kind of preliminary processing and preparation of data should be carried out when working with tasks on text data.
7. What tools and models are suitable for working with text data?
8. Describe what kind of preliminary processing and preparation of data needs to be carried out when working with tasks on tabular data.
9. What tools and models are suitable for working with tabular data?
10. What are time series and what is the peculiarity of working with time series prediction tasks?
11. What machine learning models are used to solve time series prediction problems?
12. What is a vector representation of text? Why is it needed? What vector models do you know?
13. What are the approaches to processing data gaps? What methods of filling in the gaps do you know? What are their features?
14. What is data normalization and scaling? Why are they needed in data analysis tasks?
15. What approaches to coding categorical features do you know? Give examples when it is acceptable to use each of the approaches.
16. What features for machine learning model learning can be extracted from the data types "timestamp" and "datetime"?
17. What features can be extracted from a time series to be able to solve the problem of a time series using a linear model?

18. What are the methods of increasing the training sample in the case of solving the problem of analyzing data on images.
19. Describe the key metrics of the quality of machine learning models in solving the classification problem. Which ones are not sensitive to unbalanced data?
20. Describe the key metrics of the quality of machine learning models in solving the regression problem
21. Describe the process of selecting features that positively affect the quality of the final machine learning model.
22. What is the purpose of evaluating the quality of the model. What approaches to model quality control do you know? What are the advantages and disadvantages of the main approaches.
23. Name the types of ensemble machine learning models.
24. Tell us the features of working with the XGBoost framework
25. Tell us the features of working with the LightGBM framework
26. Tell us the features of working with the CatBoost framework
27. Describe the structure of the time series.
28. 28. What time series forecasting tools do you know?
29. 29. Describe the process of preprocessing text information.
30. 30. Tell us what tf-idf is.
- 31.
32. Describe the device of an artificial neuron
33. Describe the learning process of an artificial neuron
34. Describe the activation functions that you know. Conduct a comparative analysis of them
35. Describe the device of a fully connected neural network.
36. What tasks can a neural network solve? How does the output layer and the activation function on the output layer change depending on the task being solved?
37. Describe the learning process of a neural network. What is the method of error back propagation based on?
38. Tell us about the quality metrics of neural networks for solving various tasks. Conduct a comparative analysis
39. Tell us about the loss functions of neural networks for solving various problems. Conduct a comparative analysis
40. What is neural network retraining? What methods of combating retraining do you know? How can the fact of retraining be revealed?
41. What hyper parameters of the neural network do you know? What do they affect? How does the process of selecting hyper parameters work?
42. What is gradient descent? Describe how gradient descent and stochastic gradient descent work.
43. Describe the necessary properties of error functions for which gradient descent can be applied.
44. What modifications of gradient descent do you know? Conduct a comparative analysis of gradient descent modifications
45. Describe the disadvantages of fully connected neural networks for solving image processing problems.
46. Describe the device of a convolutional neural network.
47. Describe how the convolutional layer works.
48. What are the kernels in the convolutional layer for? How are they selected?
49. How is the process of transferring learning in convolutional neural networks?
50. What modern neural network architectures for solving image processing problems do you know? Make a brief comparative analysis.
51. Describe the principle of operation of a neural network with ResNet architecture

52. Describe the principle of operation of a neural network with the Inception architecture
53. Describe the principle of operation of a neural network with the Xception architecture
54. Describe the principle of operation of a neural network with the VGG-16/19 architecture
55. Describe the principle of operation of recurrent neural networks? What tasks do they solve?
56. Describe the principle of operation of the LSTM block in a neural network.
57. Describe the principle of operation of the GRU block in a neural network.
58. Describe the principle of operation of bidirectional recurrent neural networks. What are they used for?
59. What tasks do neural networks solve when working with images?
60. What tasks do neural networks solve when working with text information?
61. What vector models of texts do you know? Conduct a comparative analysis of them
62. What is word2vec? Describe the process of its training and application features.