

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Код модуля
М.4.2

Модуль
Спортивный анализ данных

Екатеринбург, 2021

Оценочные материалы по модулю составлены авторами:

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

1. СТРУКТУРА И ОБЪЕМ МОДУЛЯ Спортивный анализ данных

№ 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.	Спортивный анализ данных	3 з.е./ 108 час.	Зачет
ИТОГО по модулю:		3 з.е./ 108 час.	

2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО МОДУЛЮ

Не предусмотрено.

Раздел 3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ 1
СПОРТИВНЫЙ АНАЛИЗ ДАННЫХ

Модуль М.4.2 Спортивный анализ данных

Оценочные материалы составлены автором(ами):

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

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

Таблица 1

Код и наименование компетенций, формируемые с участием дисциплины	Планируемые результаты обучения (индикаторы)	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	3	3
<p>УК-1 - Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий, в том числе в цифровой среде</p>	<p>З-2 - Определять этапы разработки стратегии действий, в том числе в цифровой среде, и методы решения проблемных ситуаций</p> <p>У-1 - Выявлять проблемные ситуации, используя методы системного подхода и критического анализа</p> <p>У-3 - Анализировать проблемную ситуацию, выявлять и определять способы ее разрешения</p> <p>П-2 - Использовать методы критического анализа и системного подхода в разработке стратегии действий для решения проблемных ситуаций, в том числе в цифровой среде</p> <p>Д-1 - Демонстрировать аналитические способности и критическое мышление</p>	<p>1) Контрольная работа (1 шт.)</p> <p>2) Домашние работы (2 шт.)</p> <p>3) Выполнение практических работ</p> <p>4) Зачет</p>

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 Спортивный анализ данных	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.	Спортивный анализ данных	18	18	0	36	Зачет	41,65	66,35	108	3

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

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: лекционным, практическим занятиям.		13,5 час.
2.	Выполнение и оформление мероприятий текущего контроля:		
2.1	Контрольная работа	1	5 час.
2.2	Домашняя работа	2	10 час.
3.	Подготовка к зачету	1	12 час.
4.	Самостоятельное изучение материала		92,2 час.
Итого на СРС по дисциплине:			132,7 час.

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

Процедуры текущей и промежуточной аттестации по дисциплине

1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.5		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Контрольная работа</i>	2 сем.	70
<i>Самостоятельное изучение материала</i>	2 сем.	30
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 0.5		
Промежуточная аттестация по лекциям – <i>Зачет</i>		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.5		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0.5		

Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Выполнение и оформление практических работ</i>	2 сем.	50
<i>Домашняя работа №1</i>	2 сем.	25
<i>Домашняя работа №2</i>	2 сем.	25
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям – 1		
Промежуточная аттестация по практическим/семинарским занятиям – не предусмотрена		
Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям – 0		
3. Лабораторные занятия: Не предусмотрены		
коэффициент значимости совокупных результатов лабораторных занятий – 0		

4. 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

Результаты обучения	Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам
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).

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

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	Обзор инструментов применяемых для анализа данных и машинного обучения. Обзор библиотеки Numpy. Работа с библиотекой Pandas (Matplotlib, seaborn)
2	Модели машинного обучения для решения задач классификации

3	Модели машинного обучения для решения задач регрессии
4	Метрики качества при решении задач классификации и регрессии
5	Предварительная обработка данных при решении задач анализа данных. Преобразование и создание новых признаков для решения задач анализа данных
6	Методы выбора лучших признаков для решения задач классификации и регрессии
7	Методы подбора гипер параметров моделей машинного обучения и построения конвейеров автоматизации решения задач анализа данных
8	Работа с временными рядами
9	Задачи обучения без учителя: понижение размерности Задачи обучения без учителя: поиск аномалий
10	Обработка естественного языка: предварительная обработка
11	Обработка естественного языка: векторные модели
12	Введение в нейронные сети. Нейронные сети для решения задач регрессии
13	Нейронные сети для решения задач классификации
14	Нейронные сети для решения задач обработки изображений
15	Нейронные сети для решения задач обработки изображений: современные архитектуры
16	Нейронные сети для решения задач обработки естественного языка
17	Нейронные сети для решения задач обработки естественного языка: современные архитектуры
18	Нейронные сети для решения задач предсказания временных рядов

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, вариант №1

Присоединитесь к соревнования на платформе Kaggle.com, проведите первичный анализ данных соревнования, выполните предварительную обработку данных и применяя различные модели машинного обучения добейтесь качества решения указанного в условии задания. Решается задача классификации объектов.

Контрольная работа, вариант №2

Присоединитесь к соревнования на платформе Kaggle.com, проведите первичный анализ данных соревнования, выполните предварительную обработку данных и применяя различные модели машинного обучения добейтесь качества решения указанного в условии задания. Решается задача регрессии.

Примерные задания в составе контрольных работ:

Перейдите по ссылке <https://www.kaggle.com/t/993ac81ad69a477580bd8cf1beca9a32>

Там вы увидите соревнование на платформе kaggle.com

На вкладке «Overview» вы увидите описание задания.

На вкладке «Data» находятся все необходимые данные, которые будут использованы в соревновании.

На вкладке «Notebooks» в разделе «Public» находится базовое решение с примером кода «BaseLine_1_5».

Скопировав базовое решение и запустив его у себя вы сможете достигнуть точности около значения 0,83.

Ваша задача изменяя любые параметры архитектуры нейронной сети или параметры обучения нейронной сети достигнуть наибольшего качества модели.

Результат своих экспериментов вы сможете наблюдать на вкладке «Leaderboard» сразу после загрузки ответа во вкладке «My Submissions».

Ваша цель достигнуть качества модели на тестовых данных 0.9 и выше. Тогда вы проходите задание

Когда используя LeaderBoard убедитесь, что результат вас устраивает, скачайте файл своего решения, который вы можете увидеть открыв свой рабочий блокнот в режиме просмотра (но не редактирования) и перейдя справа во вкладку «Output».

Скачанный файл в формате «.csv» загрузите сюда, на платформу курса.

5.1.5. Домашняя работа

Примерные варианты домашних работ:

Домашняя работа, вариант №1:

Дан colab notebook с заданиями по библиотеке numpy. Выполните все задания и отправьте ссылку на свое решение в google forms

Примерные тексты заданий:

1. Создайте одномерный массив numpy из 10 элементов заполненный числом 1.5
2. Создайте двумерный массив numpy из 25 элементов заполненный числом 0
3. Создайте вектор длиной 15 со случайными целыми числами в диапазоне от -10 до 10. Сделайте реверс вектора
4. Создайте матрицу 5 на 5 со случайными целыми числами в диапазоне от -15 до 15. Замените отрицательные элементы на их квадраты
5. Даны две матрицы. Убедитесь, что их можно перемножить. Произведите операцию умножения матриц, если это возможно, иначе выведите ошибку
6. Дан вектор, проверить есть в нем Nan. Проверить есть в нем знак бесконечности. Заполнить их нулями

Домашняя работа, вариант №2:

Дан colab notebook с заданиями по библиотеке pandas. Выполните все задания и отправьте ссылку на свое решение в google forms

Примерные тексты заданий:

Дан набор данных с демографической информацией о людях получающих больше 50 тысяч долларов в год и меньше. Проанализируйте набор и ответьте на вопросы используя библиотеку pandas.

1. Сколько мужчин и женщин (признак sex) представлено в этом наборе данных?
2. Каков средний возраст (признак age) женщин?
3. Какова доля граждан Германии (признак native-country)?
4. Правда ли, что люди, которые получают больше 50k, имеют как минимум высшее образование? (признак education – Bachelors, Prof-school, Assoc-acdm, Assoc-voc, Masters или Doctorate)

5. Среди кого больше доля зарабатывающих много (>50К): среди женатых или холостых мужчин (признак marital-status)? Женатыми считаем тех, у кого marital-status начинается с Married (Married-civ-spouse, Married-spouse-absent или Married-AF-spouse), остальных считаем холостыми.
6. Какое максимальное число часов человек работает в неделю (признак hours-per-week)? Сколько людей работают такое количество часов и каков среди них процент зарабатывающих много?

Домашняя работа, вариант №3:

Дан colab notebook с заданиями по библиотекам pandas, matplotlib, seaborn. Выполните все задания и отправьте ссылку на свое решение в google forms

Примерные тексты заданий:

Дан набор данных с информацией об олимпийских играх за разные годы. Проведите первичный визуальный анализ.

1. Постройте графики: линейные, столбчатые, круговые, гистограммы, ящики с усами - по каждому столбцу данных
2. Постройте точечные графики отражающие взаимосвязь между различными признаками
3. Проведите корреляционный анализ и визуализируйте его с помощью тепловой карты
4. Сделайте краткое описание к каждому графику, которое отразит информацию, которую вы увидели на визуализации.

Домашняя работа, вариант №4:

Дан colab notebook с заданиями по библиотекам sklearn. Выполните все задания и отправьте ссылку на свое решение в google forms

Примерные тексты заданий:

Дан набор данных с информацией о стоимостях квартир в зависимости от их параметров. Выполните следующие задания.

1. Проведите первичный анализ: статистический и визуальный и презентуйте его.
2. Проведите предварительную обработку данных: заполнение пропусков, обработка выбросов, генерация новых признаков.
3. Примените различные модели машинного обучения и добейтесь указанного в задании качества модели по заданной метрике.

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:

1. Назовите ключевые типы данных, которые могут быть в задачах интеллектуального анализа данных.
2. Назовите задачи, которые могут решаться при обучении моделей машинного обучения с учителем. Приведите примеры таких задач.
3. Назовите основные этапы решения задачи интеллектуального анализа данных.
4. Опишите, какую предварительную обработку и подготовку данных необходимо провести при работе с задачами на изображениях.
5. Какие инструменты и модели подходят для работы с изображениями?
6. Опишите, какую предварительную обработку и подготовку данных необходимо провести при работе с задачами на текстовых данных.
7. Какие инструменты и модели подходят для работы с текстовыми данными?
8. Опишите, какую предварительную обработку и подготовку данных необходимо провести при работе с задачами на табличных данных.
9. Какие инструменты и модели подходят для работы с табличными данными?
10. Что из себя представляют временные ряды и в чем особенность работы с задачами на предсказание временных рядов?
11. Какие модели машинного обучения применяются для решения задач на предсказание временных рядов?
12. Что такое векторное представление текста? Зачем оно нужно? Какие векторные модели вы знаете?
13. Назовите подходы к обработке пропусков в данных. Какие методики заполнения пропусков вы знаете? Какие у них особенности
14. Что такое нормализация и шкалирование данных? Для чего они нужны в задачах анализа данных?
15. Какие подходы к кодированию категориальных признаков вы знаете? Приведите примеры, когда приемлемо использовать каждый из подходов.
16. Какие признаки для обучения модели машинного обучения можно извлечь из типов данных «timestamp» и «datetime»?
17. Какие признаки можно извлечь из временного ряда, чтобы получить возможность решить задачу временного ряда посредством линейной модели?
18. Назовите методы увеличения обучающей выборки в случае решения задачи анализа данных на изображениях.
19. Опишите ключевые метрики качества моделей машинного обучения при решении задачи классификации. Какие из них не чувствительны к несбалансированным данным?
20. Опишите ключевые метрики качества моделей машинного обучения при решении задачи регрессии
21. Опишите процесс выбора признаков, которые положительно влияют на качество итоговой модели машинного обучения.
22. Для чего нужна оценка качества модели. Какие подходы к проверке качества моделей вы знаете? Назовите преимущества и недостатки основных подходов.
23. Назовите типы ансамблевых моделей машинного обучения.
24. Расскажите особенности работы с фреймворком XGBoost

25. Расскажите особенности работы с фреймворком LightGBM
26. Расскажите особенности работы с фреймворком CatBoost
27. Опишите устройство временного ряда.
28. Какие инструменты для прогнозирования временного ряда вы знаете?
29. Опишите процесс предварительной обработки текстовой информации.
30. Расскажите что такое tf-idf.
31. Опишите устройство искусственного нейрона
32. Опишите процесс обучения искусственного нейрона
33. Опишите функции активации, которые вы знаете. Проведите их сравнительный анализ
34. Опишите устройство полносвязной нейронной сети.
35. Какие задачи может решать нейронная сеть? Как изменяется выходной слой и функция активации на выходном слое в зависимости от решаемой задачи?
36. Опишите процесс обучения нейронной сети. На чем основан метод обратного распространения ошибки?
37. Расскажите про метрики качества нейронных сетей для решения различных задач. Проведите сравнительный анализ
38. Расскажите про функции потерь нейронных сетей для решения различных задач. Проведите сравнительный анализ
39. Что такое переобучение нейронной сети? Какие методы борьбы с переобучением вы знаете? Как можно выявить факт переобучения?
40. Какие гипер параметры нейронной сети вы знаете? На что они влияют? Как происходит процесс подбора гипер параметров?
41. Что такое градиентный спуск? Опишите как работает градиентный спуск и стохастический градиентный спуск.
42. Опишите необходимые свойства функций ошибок, для которых можно применить градиентный спуск.
43. Какие модификации градиентного спуска вы знаете? Проведите сравнительный анализ модификаций градиентного спуска
44. Опишите недостатками полносвязных нейронных сетей для решения задач обработки изображений.
45. Опишите устройство сверточной нейронной сети.
46. Опишите принцип работы сверточного слоя.
47. Для чего нужны ядра в сверточном слое? Как они подбираются?
48. Как происходит процесс переноса обучения в сверточных нейронных сетях?
49. Какие современные архитектуры нейронных сетей для решения задач обработки изображений вы знаете? Проведите краткий сравнительный анализ.
50. Опишите принцип работы нейронной сети с архитектурой ResNet
51. Опишите принцип работы нейронной сети с архитектурой Inception
52. Опишите принцип работы нейронной сети с архитектурой Xception
53. Опишите принцип работы нейронной сети с архитектурой VGG-16/19
54. Опишите принцип работы рекуррентных нейронных сетей? Какие задачи они решают?
55. Опишите принцип работы LSTM блока в нейронной сети.
56. Опишите принцип работы GRU блока в нейронной сети.
57. Опишите принцип работы двунаправленных рекуррентных нейронных сетей. Для чего они применяются?
58. Какие задачи решают нейронные сети при работе с изображениями?
59. Какие задачи решают нейронные сети при работе с текстовой информацией?
60. Какие векторные модели текстов вы знаете? Проведите их сравнительный анализ
61. Что такое word2vec? Опишите процесс его обучения и особенности применения.