

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
2021 г.



РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля
М.4.2

Модуль
Спортивный анализ данных

Екатеринбург, 2021

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа Прикладной анализ данных	Код ОП 09.04.02
Направление подготовки Информационные системы и технологии	Код направления и уровня подготовки 09.04.02

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

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

Программа модуля составлена 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

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

Протокол № 4 от 28.05.2021 г.

Согласовано:

Управление образовательных программ



Р.Х.Токарева

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Спортивный анализ данных

1.1. Аннотация содержания модуля

Модуль «Спортивный анализ данных» состоит из одноименной дисциплины.

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.	Спортивный анализ данных	3 з.е. / 108 ч.
ИТОГО по модулю:		3 з.е. / 108 ч.

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

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения (индикаторы)
1	2	3
Спортивный анализ данных	УК-1 - Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий, в том числе в цифровой среде	З-2 - Определять этапы разработки стратегии действий, в том числе в цифровой среде, и методы решения проблемных ситуаций У-1 - Выявлять проблемные ситуации, используя методы системного подхода и критического анализа У-3 - Анализировать проблемную ситуацию, выявлять и определять способы ее разрешения

		<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>
--	--	--

1.5. Форма обучения

Обучение по дисциплине модуля может осуществляться в очной форме.

2. СОДЕРЖАНИЕ И ОБЕСПЕЧЕНИЕ РЕАЛИЗАЦИИ ДИСЦИПЛИН МОДУЛЯ

**ПРОГРАММА МОДУЛЯ
СПОРТИВНЫЙ АНАЛИЗ ДАННЫХ**

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

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
СПОРТИВНЫЙ АНАЛИЗ ДАННЫХ**

Рабочая программа дисциплины составлена 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

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

Протокол № 4 от 28.05.2021 г.

2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ Спортивный анализ данных

2.1. Технологии обучения, используемые при изучении дисциплины модуля

– Традиционная (репродуктивная) технология.

2.2. Содержание дисциплины

Таблица 3

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Обзор инструментов применяемых для анализа данных и машинного обучения	Знакомство с инструментами: jupyter notebook, google colaboratory, 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 признаков для решения задач классификации и регрессии	Методы отбора лучших для решения поставленной 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 на государственном языке Российской Федерации.

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>.
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>.
3. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. — 2-е изд., испр. — М.: ДМК Пресс, 2018. — 652 с. . — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/107901>.
4. 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>.

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://lib.urfu.ru/>
6. Портал информационно-образовательных ресурсов УрФУ <https://study.urfu.ru/>
7. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
8. Университетская библиотека ONLINE – <https://biblioclub.ru/>
9. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks)
<http://www.bibliocomplectator.ru/available>
10. Электронные информационные ресурсы Российской государственной библиотеки
<https://www.rsl.ru/>
11. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.ru/>

2.5. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Спортивный анализ данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 4

№ 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 – https://www.python.org/