

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
«__» _____

РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1159380	Спортивный анализ данных

Екатеринбург

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа 1. Разработка программно-информационных систем 2. Разработка и управление в программных проектах	Код ОП 1. 09.04.04/33.01 2. 09.04.04/33.02
Направление подготовки 1. Программная инженерия	Код направления и уровня подготовки 1. 09.04.04

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Созыкин Андрей Владимирович	кандидат технических наук, без ученого звания	Доцент	информационных технологий и систем управления

Согласовано:

Управление образовательных программ

Р.Х. Токарева

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Спортивный анализ данных

1.1. Аннотация содержания модуля

Модуль состоит из дисциплины «Семинар по спортивному анализу данных». Содержание данной дисциплины модуля позволяет студентам изучить современные методы и инструменты анализа данных. Рассматриваются различные формы представления данных: табличные, текстовые, а также изображения. Дисциплина призвана развить творческий подход в решении задач извлечения интеллектуальных данных для решения поставленных задач. При решении задач используются передовые методы, такие как, машинное обучение и глубокие нейронные сети. Большому прогрессу и расширению кругозора способствует подход к обучению на основе соревновательной платформы. Студенты могут не только решать свои задачи на основе полученного материала, но и наблюдать передовые методы и технологии решения подобных задач участниками соревнований со всего мира.

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
1	Семинар по спортивному анализу данных	9
ИТОГО по модулю:		9

1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	Не предусмотрены
Постреквизиты и кореквизиты модуля	Не предусмотрены

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
Семинар по спортивному анализу данных	УК-1 - Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий, в	У-1 - Выявлять проблемные ситуации, используя методы системного подхода и критического анализа У-3 - Анализировать проблемную ситуацию, выявлять и определять способы ее разрешения

<p>том числе в цифровой среде</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>УК-4 - Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия</p>	<p>У-2 - Воспринимать и анализировать содержание письменных и устных текстов на родном и иностранном (ых) языках с целью определения значимой информации</p>
<p>ПК-6 - Способен модернизировать программный продукт и его окружение</p> <p>(Разработка и управление в программных проектах)</p>	<p>П-1 - Осуществлять анализ проблем и изменений программного средства в IT-проектах</p>
<p>ПК-7 - Способен самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности</p> <p>(Разработка и управление в программных проектах)</p>	<p>З-3 - Перечислить современные методы и инструменты анализа данных</p> <p>У-3 - Выбирать, анализировать и обобщать научно-техническую информацию по тематике исследования</p> <p>П-2 - Осуществлять подготовку к защите научного исследования</p> <p>П-3 - Иметь практический опыт использования творческого подхода в решении задач извлечения интеллектуальных данных</p>
<p>ПК-7 - Способен самостоятельно приобретать с помощью информационных технологий и</p>	<p>З-3 - Перечислить современные методы и инструменты анализа данных</p>

	<p>использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности</p> <p>(Разработка программно-информационных систем)</p>	<p>У-3 - Выбирать, анализировать и обобщать научно-техническую информацию по тематике исследования</p> <p>П-2 - Осуществлять подготовку к защите научного исследования</p> <p>П-3 - Иметь практический опыт использования творческого подхода в решении задач извлечения интеллектуальных данных</p>
--	---	--

1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной формах.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Семинар по спортивному анализу данных

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Созыкин Андрей Владимирович	кандидат технических наук, без ученого звания	Доцент	информационных технологий и систем управления

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

Протокол № 2 от 07.03.2019 г.

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Созыкин Андрей Владимирович, Доцент, информационных технологий и систем управления

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Базовый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.

1.2. Содержание дисциплины

Таблица 1.1

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 операции 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 гиперпараметров моделей машинного обучения и построения конвейеров автоматизации решения задач анализа данных	Инструменты построения конвейеров обработки данных в машинном обучении. Методы и инструменты автоматического подбора гиперпараметров 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

1.3. Направление, виды воспитательной деятельности и используемые технологии

Направления воспитательной деятельности сопрягаются со всеми результатами обучения компетенций по образовательной программе, их освоение обеспечивается содержанием всех дисциплин модулей.

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Семинар по спортивному анализу данных

Электронные ресурсы (издания)

1. Гласс, Гласс Р., Овчинников, В.; Программирование и конфликты 2.0. Теория и практика программной инженерии; Символ-Плюс, Санкт-Петербург ; Москва; 2010 (1 экз.)
2. Ехлаков, Ю. П.; Введение в программную инженерию : учебное пособие.; Томский государственный университет систем управления и радиоэлектроники, Эль Контент, Томск; 2011; <http://www.iprbookshop.ru/13923.html> (Электронное издание)
3. Липаев, В. В.; Программная инженерия сложных заказных программных продуктов : учебное пособие.; МАКС Пресс, Москва; 2014; <http://www.iprbookshop.ru/27297.html> (Электронное издание)
4. , Фролова, Е. А.; Методические указания по дисциплине Программная инженерия; Московский технический университет связи и информатики, Москва; 2013; <http://www.iprbookshop.ru/61752.html> (Электронное издание)
5. Сараев, П. В.; Методы машинного обучения : методические указания и задания к лабораторным работам по курсу.; Липецкий государственный технический университет, ЭБС АСВ, Липецк; 2017; <http://www.iprbookshop.ru/83183.html> (Электронное издание)
6. Пролубников, А. В.; Математические методы распознавания образов : учебное пособие.; Издательство Омского государственного университета, Омск; 2020; <http://www.iprbookshop.ru/108119.html> (Электронное издание)
7. , Боровиков, В. П.; Нейронные сети. STATISTICA Neural Networks. Методология и технологии современного анализа данных; Горячая линия - Телеком, Москва; 2008 (1 экз.)

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/>
4. Электронные информационные ресурсы Российской государственной библиотеки <https://www.rsl.ru/>
5. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.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

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Семинар по спортивному анализу данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Персональные компьютеры по количеству обучающихся Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет Мультимедийный проектор с экраном.	Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES Adobe Acrobat Professional 2017 Multiple Platforms Используется бесплатно-распространяемое программное обеспечение: 1. Python – https://www.python.org/