

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

<b>Код модуля</b>	<b>Модуль</b>
1161465	Прикладной анализ данных

Екатеринбург

<b>Перечень сведений о рабочей программе модуля</b>	<b>Учетные данные</b>
<b>Образовательная программа</b> 1. Управление техносферными рисками	<b>Код ОП</b> 1. 20.04.01/33.01
<b>Направление подготовки</b> 1. Техносферная безопасность	<b>Код направления и уровня подготовки</b> 1. 20.04.01

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	Комоцкий Евгений Игоревич	без ученой степени, без ученого звания	Ассистент	Базовая кафедра "Аналитика больших данных и методы видеопроцессинга"
2	Федотов Олег Васильевич	без ученой степени, без ученого звания	Старший преподаватель	интеллектуальных информационных технологий

**Согласовано:**

Управление образовательных программ

Р.Х. Токарева

# 1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Прикладной анализ данных

## 1.1. Аннотация содержания модуля

Модуль “Прикладной анализ данных” состоит из одноименной дисциплины и позволяет магистранту приобрести теоретические знания и практические навыки по анализу данных с использованием средств языка программирования Python, специальных библиотек для анализа данных, а также дает широкий обзор статистических методов и инструментов анализа и визуализации данных.

## 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	Прикладной анализ данных	3
ИТОГО по модулю:		3

## 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
Прикладной анализ данных	ПК-6 - Способен использовать цифровые решения для работы с потоками ESG-данных и составления отчетности организации	З-1 - Описать современные информационно-коммуникативные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач З-2 - Описывать современные техники и методики сбора данных для решения управленческих и исследовательских задач З-4 - Сделать обзор основных методов моделирования и математического анализа,

		<p>применимых для формализации и решения задач в области техносферной безопасности</p> <p>З-5 - Описать технологии хранения и обработки больших данных в организации: базы данных, хранилища данных, распределенную и параллельную обработку данных, вычисления в оперативной памяти</p> <p>У-2 - Выбирать оптимальные методы обработки и анализа данных, в том числе используя интеллектуальные информационно-аналитические системы при решении управленческих и исследовательских задач</p> <p>У-6 - Выбирать оптимальные современные методы и инструментальные средства анализа больших данных, методы интерпретации и визуализации больших данных</p> <p>П-2 - Создавать отчеты, рефераты, статьи, оформленные в соответствии с предъявляемыми требованиями с использованием интеллектуальных информационно-аналитических систем</p>
--	--	--

### 1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной формах.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Прикладной 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	Комоцкий Евгений Игоревич	без ученой степени, без ученого звания	Ассистент	Базовая кафедра "Аналитика больших данных и методы видеоанализа"
2	Федотов Олег Васильевич	без ученой степени, без ученого звания	Старший преподаватель	интеллектуальных информационных технологий

**Рекомендовано учебно-методическим советом института Фундаментального образования**

Протокол № 5 от 25.05.2023 г.

# 1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- **Комоцкий Евгений Игоревич, Ассистент, Базовая кафедра ”Аналитика больших данных и методы видеоанализа”**
- **Федотов Олег Васильевич, Старший преподаватель, интеллектуальных информационных технологий**

## 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
P1	Обзор программных продуктов для анализа данных	Цели и задачи анализа данных. Особенности программ Vortex, SPSS, Statistica. Анализ данных с использованием средств языка программирования Python, R.
P2	Первичная обработка данных	Классы задач обработки данных. Обработка и анализ данных. Проблемы, возникающие при постановке задачи подбора методов обработки данных. Основы статистики.
P3	Основы языка программирования Python	Введение в синтаксис языка. Jupyter Notebook. Google Colab. Библиотеки numpy, pandas.
P4	Введение в машинное обучение	Введение в машинное обучение. Основные определения. Этапы разработки моделей машинного обучения. Постановки задач машинного обучения. Примеры прикладных задач и их типы: классификация, регрессия, ранжирование, кластеризация, поиск структуры в данных.
P5	Введение в платформу Kaggle	Регистрация на платформе. Интерфейс платформы. Образовательные курсы. Соревнования. Форум.
P6	Оценка качества алгоритмов	Регрессия: квадратичные и абсолютные потери, абсолютные логарифмические отклонения. Примеры использования. Классификация: доля верных ответов, ее недостатки. Точность и полнота, их объединение: арифметическое среднее, минимум, гармоническое среднее (F-мера). Практические

		особенности кросс-валидации. Стратификация. Потенциальные проблемы с разбиением зависимой или динамической выборки.
<b>P7</b>	Композиции алгоритмов	Простейший пример: уменьшение дисперсии при усреднении алгоритмов методом бутстреп. Блендинг алгоритмов. Понятие смещения и разброса (иллюстрация на примере линейных методов и решающих деревьев). Уменьшение разброса с помощью усреднения. Случайный лес. Оценка out-of-bag.
<b>P8</b>	Искусственные нейронные сети	Введение в deep learning. Обучение глубоких нейронных сетей. Библиотека Keras.
<b>P9</b>	Построение отчетов по результатам анализа больших данных	Построение интерактивных отчетов (dashboards). Способы визуализации данных. Сравнительный анализ современных систем Business Intelligence, Tableau; QlikView; Prognoz Platform; Power BI.

### 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

### Прикладной анализ данных

#### Электронные ресурсы (издания)

1. Комлева, Н. В.; Структуры и алгоритмы компьютерной обработки данных : учебное пособие.; Московский государственный университет экономики, статистики и информатики, Москва; 2004; <https://biblioclub.ru/index.php?page=book&id=93226> (Электронное издание)
2. Мейер, Б., Б.; Инструменты, алгоритмы и структуры данных; Национальный Открытый Университет «ИНТУИТ», Москва; 2016; <https://biblioclub.ru/index.php?page=book&id=429033> (Электронное издание)

#### Печатные издания

1. Мазуров, В. Д.; Метод комитетов в задачах оптимизации и классификации; Наука, Москва; 1990 (6 экз.)
2. Редько, В. Г., Малинецкий, Г. Г.; Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики; URSS, Москва; 2015 (2 экз.)
3. Редько, В. Г.; Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики; [КомКнига], Москва; 2005 (1 экз.)
4. Котов, В. М.; Алгоритмы и структуры данных : [учебное пособие для вузов].; БГУ, Минск; 2011 (1 экз.)

#### Профессиональные базы данных, информационно-справочные системы

1. Электронный научный архив УрФУ <https://elar.urfu.ru/>
2. Зональная научная библиотека (УрФУ) - <http://lib.urfu.ru/>

3. Портал информационно-образовательных ресурсов УрФУ <https://study.urfu.ru/>
4. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
5. Университетская библиотека ONLINE – <https://biblioclub.ru/>
6. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks)
7. <http://www.bibliocomplectator.ru/available>
8. Электронные информационные ресурсы Российской государственной библиотеки  
<https://www.rsl.ru/>
9. 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. 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. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
5. eBook Collections Springer Nature – <https://link.springer.com/>

## **3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Прикладной анализ данных**

**Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением**

Таблица 3.1

<b>№ п/п</b>	<b>Виды занятий</b>	<b>Оснащённость специальных помещений и помещений для самостоятельной работы</b>	<b>Перечень лицензионного программного обеспечения</b>
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Подключение к сети Интернет	Office Professional 2003 Win32 Russian CD-ROM Бесплатно распространяемое программное обеспечение: Python – <a href="https://www.python.org/">https://www.python.org/</a> Веб - среда разработки для языка



			<p>программирования Python: google colab -  <a href="https://colab.research.google.com/">https://colab.research.google.com/</a>          Anaconda solution -  <a href="https://www.anaconda.com/">https://www.anaconda.com/</a>          Система распределенной обработки данных Apache Spark –  <a href="https://spark.apache.org/">https://spark.apache.org/</a>          Кластер HDInsight от Microsoft Azure</p>
2	Практические занятия	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Периферийное устройство</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>Office Professional 2003 Win32 Russian CD-ROM</p> <p>Бесплатно</p> <p>распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a>          Веб - среда разработки для языка программирования Python: google colab -  <a href="https://colab.research.google.com/">https://colab.research.google.com/</a>          Anaconda solution -  <a href="https://www.anaconda.com/">https://www.anaconda.com/</a>          Система распределенной обработки данных Apache Spark –  <a href="https://spark.apache.org/">https://spark.apache.org/</a>          Кластер HDInsight от Microsoft Azure</p>
3	Лабораторные занятия	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Периферийное устройство</p>	<p>Office Professional 2003 Win32 Russian CD-ROM</p> <p>Бесплатно</p> <p>распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a></p>

		<p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>Веб - среда разработки для языка программирования Python: google colab -</p> <p><a href="https://colab.research.google.com/">https://colab.research.google.com/</a></p> <p>Anaconda solution -</p> <p><a href="https://www.anaconda.com/">https://www.anaconda.com/</a></p> <p>Система распределенной обработки данных Apache Spark –</p> <p><a href="https://spark.apache.org/">https://spark.apache.org/</a></p> <p>Кластер HDInsight от Microsoft Azure</p>
4	Консультации	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Подключение к сети Интернет</p>	<p>Office Professional 2003 Win32 Russian CD-ROM</p>
5	Текущий контроль и промежуточная аттестация	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>Office Professional 2003 Win32 Russian CD-ROM</p> <p>Бесплатно</p> <p>распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a></p> <p>Веб - среда разработки для языка программирования Python: google colab -</p> <p><a href="https://colab.research.google.com/">https://colab.research.google.com/</a></p> <p>Anaconda solution -</p> <p><a href="https://www.anaconda.com/">https://www.anaconda.com/</a></p> <p>Система распределенной обработки данных Apache Spark –</p> <p><a href="https://spark.apache.org/">https://spark.apache.org/</a></p> <p>Кластер HDInsight от Microsoft Azure</p>

6	Самостоятельная работа студентов	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Подключение к сети Интернет</p>	<p>Office Professional 2003 Win32 Russian CD-ROM</p> <p>Бесплатно распространяемое программное обеспечение:</p> <p>Python – <a href="https://www.python.org/">https://www.python.org/</a></p> <p>Веб - среда разработки для языка программирования Python: google colab - <a href="https://colab.research.google.com/">https://colab.research.google.com/</a></p> <p>Anaconda solution - <a href="https://www.anaconda.com/">https://www.anaconda.com/</a></p> <p>Система распределенной обработки данных Apache Spark – <a href="https://spark.apache.org/">https://spark.apache.org/</a></p> <p>Кластер HDInsight от Microsoft Azure</p>
---	----------------------------------	---	--