



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Магнитогорский государственный технический университет им. Г.И. Носова»



УТВЕРЖДАЮ  
Директор ИЭиАС  
В.Р. Храмшин

03.03.2021 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

***БОЛЬШИЕ ДАННЫЕ В ОБРАЗОВАНИИ***

Направление подготовки (специальность)  
44.04.01 Педагогическое образование

Направленность (профиль/специализация) программы  
Информационные технологии в образовании

Уровень высшего образования - магистратура

Форма обучения  
очная

Институт/ факультет	Институт энергетики и автоматизированных систем
Кафедра	Бизнес-информатики и информационных технологий
Курс	1
Семестр	1

Магнитогорск  
2021 год

Рабочая программа составлена на основе ФГОС ВО - магистратура по направлению подготовки 44.04.01 Педагогическое образование (приказ Минобрнауки России от 22.02.2018 г. № 126)

Рабочая программа рассмотрена и одобрена на заседании кафедры Бизнес-информатики и информационных технологий  
18.02.2021, протокол № 6

Зав. кафедрой  Г.Н. Чусавитина


Рабочая программа одобрена методической комиссией ИЭиАС  
03.03.2021 г. протокол № 5

Председатель  В.Р. Храмшин

Рабочая программа составлена:  
доцент кафедры БИиИТ, канд. пед. наук

 Е.В. Карманова

доцент кафедры БИиИТ, д-р пед. наук

 Т.Б. Новикова

Рецензент:

Руководитель аналитических работ в ИТ-проектах ЗАО «КОНСОМ СКС», канд. пед. наук  
В.А. Ошурков



## Лист актуализации рабочей программы

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2022 - 2023 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ Г.Н. Чусавитина

---

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2023 - 2024 учебном году на заседании кафедры Бизнес-информатики и информационных

Протокол от \_\_\_\_\_ 20\_\_ г. № \_\_\_\_  
Зав. кафедрой \_\_\_\_\_ Г.Н. Чусавитина

### **1 Цели освоения дисциплины (модуля)**

Целями освоения дисциплины «Большие данные в образовании» является теоретическая и практическая подготовка студентов к работе с большими данными. Знания, полученные в результате освоения дисциплины, помогут при сборе и анализе огромных объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач практической и научно-исследовательской деятельности.

### **2 Место дисциплины (модуля) в структуре образовательной программы**

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

- Производственная - преддипломная практика
- Производственная - научно-исследовательская работа
- Выполнение и защита выпускной квалификационной работы
- Производственная - технологическая (проектно-технологическая) практика
- Оценка качества обучения и ресурсов цифровой образовательной среды
- Учебная - технологическая (проектно-технологическая) практика
- Производственная - педагогическая практика
- Компьютерное моделирование психолого-педагогических исследований
- Учебная - научно-исследовательская работа
- Теория и методика обучения информатике в системе непрерывного образования
- Облачные и мобильные технологии в образовании
- Методология и методы научного исследования
- Виртуальная и дополненная реальность в образовании

### **3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения**

В результате освоения дисциплины (модуля) «Большие данные в образовании» обучающийся должен обладать следующими компетенциями:

Код индикатора	Индикатор достижения компетенции
ПК-1	Способен анализировать результаты научных исследований, применять их при решении конкретных научно-исследовательских задач в сфере применения ИКТ в образовании, самостоятельно осуществлять научное исследование
ПК-1.1	Анализирует, систематизирует и обобщает результаты научных и научно-методических исследований в сфере применения ИКТ в образовании
ПК-1.2	Самостоятельно организывает и проводит научно-исследовательскую деятельность и использует ее результаты при решении профессиональных задач
ПК-1.3	Организует исследовательскую и проектную деятельности обучающихся

ПК-2 Способен реализовывать образовательный процесс с использованием современных образовательных технологий в цифровой образовательной среде	
ПК-2.1	Использует современные информационные технологии в педагогической деятельности (обучении и в управлении) в системе общего и дополнительного образования
ПК-2.2	Осуществляет преподавание основных и дополнительных образовательных программ с использованием дистанционных образовательных технологий и электронного обучения
ПК-2.3	Разрабатывает цифровые образовательные ресурсы, осуществляет контроль качества их создания и применения; обеспечивает безопасную работу в цифровой образовательной среде

#### 4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 5 зачетных единиц 180 акад. часов, в том числе:

- контактная работа – 76,1 акад. часов;
- аудиторная – 72 акад. часов;
- внеаудиторная – 4,1 акад. часов;
- самостоятельная работа – 68,2 акад. часов;
- в форме практической подготовки – 0 акад. час;
- подготовка к экзамену – 35,7 акад. час

Форма аттестации - экзамен

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в акад. часах)			Самостоятельная работа студента	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		Лек.	лаб. зан.	практ. зан.				
1. Введение в анализ больших данных. Предобработка								
1.1 Жизненный цикл аналитики данных. Методы анализа больших данных. Обзор программных средств для анализа больших данных	1	6	8/2И		3,9	Выполнение лабораторных заданий	Отчет по лабораторной работе	ПК-1.1, ПК-1.2
1.2 Методы поиска и сбора больших данных. Web-scraping. Технологии хранения		4	4/1И		6	Выполнение лабораторных заданий	Отчет по лабораторной работе	ПК-1.2
1.3 Методы обработки пропусков и дубликатов данных. Приведение типов данных. Технологии парсинга		4	6/1,9И		8	Выполнение лабораторных заданий	Отчет по лабораторной работе	ПК-1.2
1.4 Категоризация данных. Лемматизация, стемминг данных		2	2/2И		10	Выполнение лабораторных заданий	Отчет по лабораторной работе	ПК-1.2
Итого по разделу		16	20/6,9И		27,9			
2. Визуальный анализ данных								
2.1 Методы объединения данных.	1	2	2		10	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	ПК-1.2
2.2 Визуализация данных. Инструменты построения графиков. Гистограммы, диаграммы (с областями, столбчатые, точечные, пузырьковые, лепестковые и др), «ящики с усами».		4	4/1И		2	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	ПК-1.1, ПК-1.2

2.3 Анализ и визуализация геоданных. Библиотека folium, интерактивная визуализация геоданных.		4	2/2И		4	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	ПК-1.2
2.4 Визуализация графов и сетевых структур. Библиотека networkx, Gephi.		4	2		4,3	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	ПК-1.2
Итого по разделу		14	10/3И		20,3			
3. Анализ больших данных социальных сетей								
3.1 Анализ сетей. Степенной закон. Меры сосредоточенности и ранжирование по сети	1	4	4/2,1И		10	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	ПК-1.1, ПК-1.2
3.2 Анализ графов сообществ социальной сети		2	2/2И		10	Выполнение лабораторных заданий, изучение литературы	Отчет по лабораторной работе	
Итого по разделу		6	6/4,1И		20			
4. Экзамен								
4.1 Экзамен	1							
Итого по разделу								
Итого за семестр		36	36/14И		68,2		экзамен	
Итого по дисциплине		36	36/14И		68,2		экзамен	

## **5 Образовательные технологии**

В ходе проведения занятий предусматриваются следующие образовательные технологии:

1. Информационная лекция – последовательное изложение материала в дисциплинарной логике, осуществляемое преимущественно вербальными средствами (монолог преподавателя).

2. Лабораторная работа – организация учебной работы с реальными материальными и информационными объектами, экспериментальная работа с аналоговыми моделями реальных объектов.

3. Технологии проблемного обучения – организация образовательного процесса, которая предполагает постановку проблемных вопросов, создание учебных проблемных ситуаций для стимулирования активной познавательной деятельности студентов.

4. Лабораторное занятие на основе кейс-метода – обучение в контексте моделируемой ситуации, воспроизводящей реальные условия научной, производственной, общественной деятельности. Обучающиеся должны проанализировать ситуацию, разобраться в сути проблем, предложить возможные решения и выбрать лучшее из них. Кейсы базируются на реальном фактическом материале или же приближены к реальной ситуации.

5. Информационно-коммуникационные образовательные технологии – организация образовательного процесса, основанная на применении специализированных программных сред и технических средств работы с информацией.

6. Формы учебных занятий с использованием информационно-коммуникационных технологий:

7. Лекция-визуализация – изложение содержания сопровождается презентацией (демонстрацией учебных материалов, представленных в различных знаковых системах, в т.ч. иллюстративных, графических, аудио- и видеоматериалов).

8. Лабораторное занятие в форме презентации – представление результатов проектной или исследовательской деятельности с использованием специализированных программных сред.

В рамках практических занятий предусматривается использование средств вычислительной техники при выполнении индивидуальных заданий. Используется существующий образовательный портал университета ([newlms.magtu.ru](http://newlms.magtu.ru)) для размещения ЭУМК по дисциплине. Текущий, промежуточный и рубежный контроль проводится на образовательном портале университета.

## **6 Учебно-методическое обеспечение самостоятельной работы обучающихся**

Представлено в приложении 1.

## **7 Оценочные средства для проведения промежуточной аттестации**

Представлены в приложении 2.

## **8 Учебно-методическое и информационное обеспечение дисциплины (модуля)**

### **а) Основная литература:**

1. Федоров, Д. Ю. Программирование на языке высокого уровня Python : учебное пособие для вузов / Д. Ю. Федоров. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2020. — 161 с. — (Высшее образование). — ISBN 978-5-534-10971-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/454100>

2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование).



— ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>

**б) Дополнительная литература:**

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450166> (дата обращения: 11.05.2020).

2. Журнал «Программные продукты и системы» [Электронный ресурс]. Научно-исследовательский институт «Центрпрограммсистем» — Режим доступа: [https://e.lanbook.com/journal/2276#journal\\_name](https://e.lanbook.com/journal/2276#journal_name) — Загл. с экрана.

**в) Методические указания:**

**г) Программное обеспечение и Интернет-ресурсы:**

**Программное обеспечение**

Наименование ПО	№ договора	Срок действия лицензии
7Zip	свободно	бессрочно
Anaconda Python	свободно	бессрочно
JetBrains PyCharm Community Edition	свободно распространяемое ПО	бессрочно
LibreOffice	свободно	бессрочно
Браузер Mozilla Firefox	свободно распространяемое ПО	бессрочно

**Профессиональные базы данных и информационные справочные системы**

Название курса	Ссылка
Электронная база периодических изданий East View Information Services, ООО «ИВИС»	<a href="https://dlib.eastview.com/">https://dlib.eastview.com/</a>
Национальная информационно-аналитическая система – Российский индекс научного цитирования	URL: <a href="https://elibrary.ru/project_risc.asp">https://elibrary.ru/project_risc.asp</a>
Российская Государственная библиотека. Каталоги	<a href="https://www.rsl.ru/ru/4readers/catalogues/">https://www.rsl.ru/ru/4readers/catalogues/</a>
Электронные ресурсы библиотеки МГТУ им. Г.И. Носова	<a href="http://magtu.ru:8085/marcweb2/Default.asp">http://magtu.ru:8085/marcweb2/Default.asp</a>
Университетская информационная система РОССИЯ	<a href="https://uisrussia.msu.ru">https://uisrussia.msu.ru</a>
Международная наукометрическая реферативная и полнотекстовая база данных научных изданий «Web of science»	<a href="http://webofscience.com">http://webofscience.com</a>
Международная реферативная и полнотекстовая справочная база данных	<a href="http://scopus.com">http://scopus.com</a>
Международная база справочных изданий по всем отраслям знаний SpringerReference	<a href="http://www.springer.com/references">http://www.springer.com/references</a>

Архив научных журналов «Национальный электронно-информационный	<a href="https://archive.neicon.ru/xmlui/">https://archive.neicon.ru/xmlui/</a>
Информационная система - Нормативные правовые акты, организационно-распорядительные документы, нормативные и методические	<a href="https://fstec.ru/normotvorcheskaya/tekhnicheskaya-zashchita-informatsii">https://fstec.ru/normotvorcheskaya/tekhnicheskaya-zashchita-informatsii</a>

### **9 Материально-техническое обеспечение дисциплины (модуля)**

Материально-техническое обеспечение дисциплины включает:

Учебные аудитории для проведения занятий лекционного типа Персональный компьютер (или ноутбук) с пакетом Office, с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Доска, мультимедийный проектор, экран. Мультимедийные презентации к лекциям, учебно-наглядные пособия

Учебные аудитории для проведения лабораторных занятий Персональные компьютеры с пакетом Office; Anaconda Python, Jupyter Notebook (anaconda), с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Браузер Mozilla.

Аудитории для самостоятельной работы: компьютерные классы; читальные залы библиотеки Персональные компьютеры с пакетом Office; Anaconda Python, Jupyter Notebook (anaconda), с выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Браузер Mozilla.

Аудитории для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации Персональные компьютеры с пакетом Office и выходом в Интернет и с доступом в электронную информационно-образовательную среду университета. Anaconda Python, Jupyter Notebook (anaconda); Браузер Mozilla.

Аудитория для хранения и профилактического обслуживания учебного оборудования Мебель для хранения и обслуживания оборудования (шкафы, столы), учебно-методические материалы, компьютеры, ноутбуки, принтеры.

### **Учебно-методическое обеспечение самостоятельной работы обучающихся**

По дисциплине «Большие данные в образовании» предусмотрена аудиторная и внеаудиторная самостоятельная работа обучающихся.

Аудиторная самостоятельная работа студентов предполагает выполнение заданий по лабораторным работам, подготовка к тестированию.

Для выполнения лабораторных работ вам будут полезны следующие ресурсы:

- Большие данные и всё о них - <https://habr.com/ru/hub/bigdata/>
- Сайт компании BasegroupLabs- <https://basegroup.ru/>
- Портал открытых данных РФ - <https://data.gov.ru/>

Предметные области для определения темы проекта:

- Работа с твитами – эмодзи
- Данные с росстата
- Данные с центробанка
- Данные по экологии
- Данные по стоимости недвижимости
- Др.

Примечание. При выборе предметной области, продумайте возможность использования геоданных, поскольку во второй половине курса потребуется их обработка.

### **Лабораторная работа №1**

Тема «Сбор и хранение данных»

#### **Задания:**

1. Выберите предметную область для реализации дальнейшего проекта:

- Работа с твитами – эмодзи
- Данные с росстата
- Данные с центробанка
- Данные по экологии
- Данные по стоимости недвижимости
- Др.

2. изучите строение HTML-страниц, правила работы GET-запросов;

3. составьте несколько запросов к сайтам и соберите данные.

4. примените методы автоматизированного веб-скрейпинга при сборе данных.

### **Лабораторная работа №2**

Тема «Предобработка данных»

#### **Задания:**

1. По выбранной предметной области необходимо подготовить датафрейм.

2. Полученный датафрейм обработать на предмет пропусков, привести типы данных в соответствие, найти и убрать дубликаты данных; привести к одному виду названия столбцов.

3. Для анализа текстовых данных примените методы стемминга и лемматизации.

### **Лабораторная работа №3**

Тема «Визуализация»

#### **Задания:**

1. Изучите методы объединения данных в Python. Реализуйте объединение данных в вашем датафрейме для решения конкретной задачи.

2. По полученным данным датафрейма постройте графики, используя методы: *hist()*, *boxplot()*, *plot()*.

3. Проанализируйте полученные результаты, какие выводы можно сделать по построенным графикам?

**Лабораторная работа №4**  
**Тема «Числовые показатели»**

**Задания:**

1. Найдите среднее, медиану, количество наблюдений в выборке и разброс их значений.
2. Какие выводы можно сделать о данных в датафрейме по полученным показателям?

**Лабораторная работа №5**  
**Тема «Геоданные»**

**Задания:**

1. Рассмотрите проблему обработки и визуализации геоданных.
2. Получите геоданные по одному ваших столбцов датафрейма.
3. Проведите группировку данных, решая некоторую задачу.
4. Постойте интерактивную карту, используя библиотеку folium.

**Лабораторная работа №6**  
**Тема «Визуализация графов»**

**Задания:**

1. Создайте полносвязный граф. Полносвязный граф — граф, где каждая вершина соединена с каждой другой.
2. С использованием библиотеки networkx постройте визуализацию данного графа.

**Лабораторная работа № 7**  
**Тема «Анализ данных социальных сетей»**

**Задания:**

Подключитесь к социальной сети Вконтакте и проанализируйте свой социальный граф

Оценочные средства для проведения промежуточной аттестации

а) Планируемые результаты обучения и оценочные средства для проведения промежуточной аттестации:

Код индикатора	Индикатор достижения компетенции	Оценочные средства
ПК-1 - способен осуществлять аналитические исследования с применением технологий искусственного интеллекта		
ПК-1.1	Выявляет требования, планирует и организует аналитические работы с использованием технологий искусственного интеллекта	<p>1. Методики анализа больших данных.                  2. Дайте характеристику BigData на мировом рынке.                  3. Охарактеризуйте BigData в России.                  4. Определите понятие DataMining.                  5. Методы поиска и сбора больших данных.                  6. Технологии хранения больших данных</p> <p><i>Практическое задание:</i></p> <p>1. Реализуйте чтение данных из таблиц Excel                  2. Реализуйте чтение данных из файла csv                  3. Получите данные с помощью GET-запроса.                  4. Проанализируйте данные с помощью библиотеки BeautifulSoup</p> <p><i>Проектное задание:</i>                  Получите данные по стоимости недвижимости отдельного города за несколько лет (не менее 3 лет). На основе полученных данных ответьте на следующие вопросы:</p> <p>1. Какова тенденция изменения стоимости недвижимости?                  2. Как стоимость недвижимости отличается внутри города, от каких показателей зависит, есть ли корреляция между отдельными показателями?                  3. Какими данными необходимо дополнить датафрейм?                  4. Визуализируйте полученные данные.</p>
ПК-1.2	Осуществляет подготовку данных и проводит аналитические исследования с использованием технологий искусственного интеллекта	<p>1. Методы анализа больших данных.                  2. Обзор программных средств для анализа больших данных.                  3. Технологии хранения больших данных.                  4. Методы и подходы к предобработке больших данных                  5. Технология парсинга больших данных.                  6. Технология веб-скрейпинга.                  7. Методы визуализации данных.                  8. Методы поиска и сбора больших данных. Web-scraping                  9. Технологии хранения больших данных                  10. Сервисы для получения больших данных.</p>

Код индикатора	Индикатор достижения компетенции	Оценочные средства
		<p>Сервис открытые данные РФ.</p> <p>11. Зарубежные сервисы больших данных.</p> <p>12. Средства получения и визуализации геоданных.</p> <p>13. Данные социальных сетей. Возможности для бизнеса.</p>
		<p><i>Практические задания:</i></p> <ol style="list-style-type: none"> <li>1. Получите датафрейм, используя API любой социальной сети.</li> <li>2. Получите статистические данные по датафрейму.</li> <li>3. Обработайте пропуски и дубликаты в датафрейме</li> <li>4. Постройте диаграмму – ящик с усами по отдельным столбцам датафрейма</li> <li>5. Составьте обзор сервисов, хранящих большие данные (не менее 10 сервисов). Заполните таблицу: название сервиса, URL, описание данных, форматы хранения, методы доступа к данным.</li> </ol>
		<p><i>Проектное задание:</i></p> <p>Получите данные по стоимости недвижимости отдельного города за несколько лет (не менее 3 лет). На основе полученных данных ответьте на следующие вопросы:</p> <ol style="list-style-type: none"> <li>1. Какова тенденция изменения стоимости недвижимости?</li> <li>2. Как стоимость недвижимости отличается внутри города, от каких показателей зависит, есть ли корреляция между отдельными показателями?</li> <li>3. Какими данными необходимо дополнить датафрейм?</li> </ol> <p>Визуализируйте полученные данные.</p>
ПК-1.3	Формирует предложения по использованию результатов аналитического исследования	<p>Не формируется</p> <p>Не формируется</p> <p>Не формируется</p>

**б) Порядок проведения промежуточной аттестации, показатели и критерии оценивания:**

Промежуточная аттестация по дисциплине «Большие данные в образовании» включает теоретические вопросы, позволяющие оценить уровень усвоения обучающимися знаний, и практические задания, выявляющие степень сформированности умений и владений, проводится в форме экзамена.

Экзамен по данной дисциплине проводится в устной форме по экзаменационным билетам, каждый из которых включает 2 теоретических вопроса и одно практическое задание.

**Показатели и критерии оценивания экзамена:**

– на оценку **«отлично»** (5 баллов) – обучающийся демонстрирует высокий уровень сформированности компетенций, всестороннее, систематическое и глубокое знание учебного материала, свободно выполняет практические задания, свободно оперирует знаниями, умениями, применяет их в ситуациях повышенной сложности.

– на оценку **«хорошо»** (4 балла) – обучающийся демонстрирует средний уровень сформированности компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.

– на оценку **«удовлетворительно»** (3 балла) – обучающийся демонстрирует пороговый уровень сформированности компетенций: в ходе контрольных мероприятий допускаются ошибки, проявляется отсутствие отдельных знаний, умений, навыков, обучающийся испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.

– на оценку **«неудовлетворительно»** (2 балла) – обучающийся демонстрирует знания не более 20% теоретического материала, допускает существенные ошибки, не может показать интеллектуальные навыки решения простых задач.

– на оценку **«неудовлетворительно»** (1 балл) – обучающийся не может показать знания на уровне воспроизведения и объяснения информации, не может показать интеллектуальные навыки решения простых задач.

## Методические по выполнению индивидуальных домашних заданий

Цель методических рекомендаций - обеспечить студенту оптимальную организацию процесса изучения дисциплины, а также выполнения различных форм самостоятельной работы.

Студентам необходимо ознакомиться: с содержанием рабочей программы дисциплины (далее - РПД), с целями и задачами дисциплины, ее связями с другими дисциплинами образовательной программы, методическими разработками по данной дисциплине, имеющимся на образовательном портале и сайте кафедры, с графиком консультаций преподавателей кафедры.

### Инструкции по работе с Anaconda

Anaconda — это дистрибутивы Python и R. Он предоставляет все необходимое для решения задач по анализу и обработке данных (с применимостью к Python).

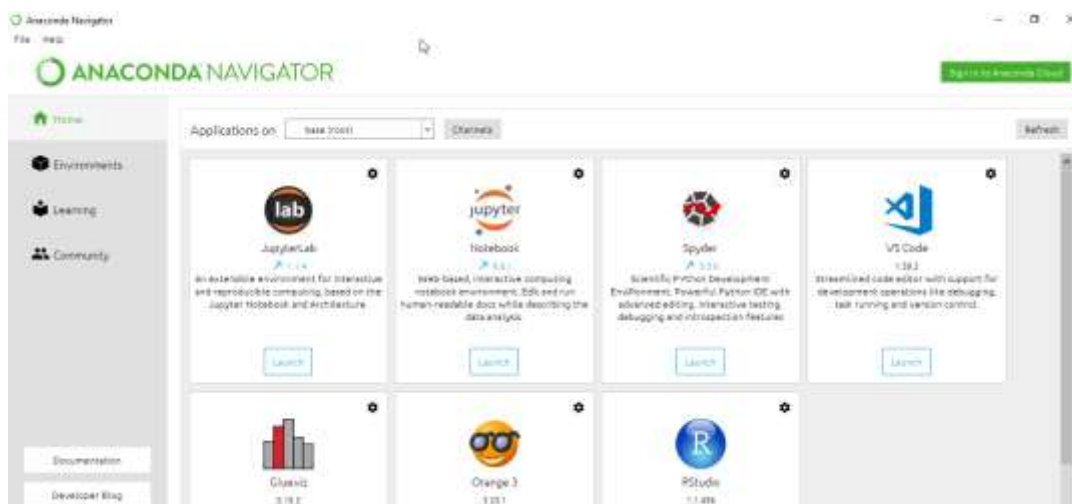
Anaconda — это набор бинарных систем, включающий в себя Scipy, Numpy, Pandas и их зависимости.

Scipy — это пакет статистического анализа.

Numpy — это пакет числовых вычислений.

Pandas — уровень абстракции данных для объединения и преобразования данных.

Anaconda Navigator — это графический интерфейс пользователя на рабочем столе (GUI), включенный в дистрибутив Anaconda, который позволяет запускать приложения и легко управлять пакетами, средами и каналами conda без использования команд командной строки. Навигатор может искать пакеты в Anaconda Cloud или в локальной репозитории Anaconda. Он доступен для Windows, MacOS и Linux.



JupyterLab — это интерактивная среда разработки для работы с блокнотами, кодом и данными.

Jupyter Notebook — удобный инструмент для создания красивых аналитических отчетов, позволяет хранить вместе код, изображения, комментарии, формулы и графики. Работа ведется в браузере.

Spyder — интерактивной IDE для научных расчетов на языке Python. Данная IDE позволяет писать, редактировать и тестировать код. Spyder предлагает просмотр и редактирование переменных с помощью GUI, динамическую интроспекцию кода, нахождение ошибок на лету и многое другое. Также, по необходимости, можно интегрировать Anaconda с другими Python IDE, включая PyCharm и Atom.

VS Code — это оптимизированный редактор кода с поддержкой таких операций разработки, как отладка, запуск задач и контроль версий.

Glueviz — используется для визуализации многомерных данных в файлах. Он исследует отношения внутри и между связанными наборами данных.

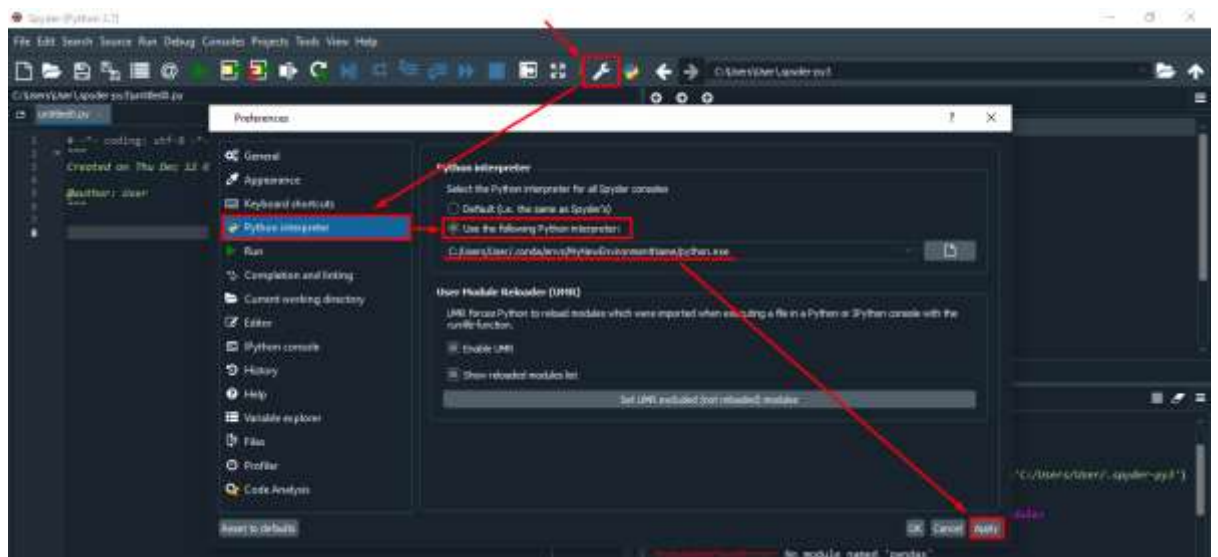
Orange 3 — это основанная на компонентах структура интеллектуального анализа данных. Это может быть использовано для визуализации данных и анализа данных. Рабочие процессы в Orange 3 очень интерактивны и предоставляют большой набор инструментов.

RStudio — это набор интегрированных инструментов, предназначенных для повышения продуктивности работы с R. Он включает в себя основы R и Notebooks.

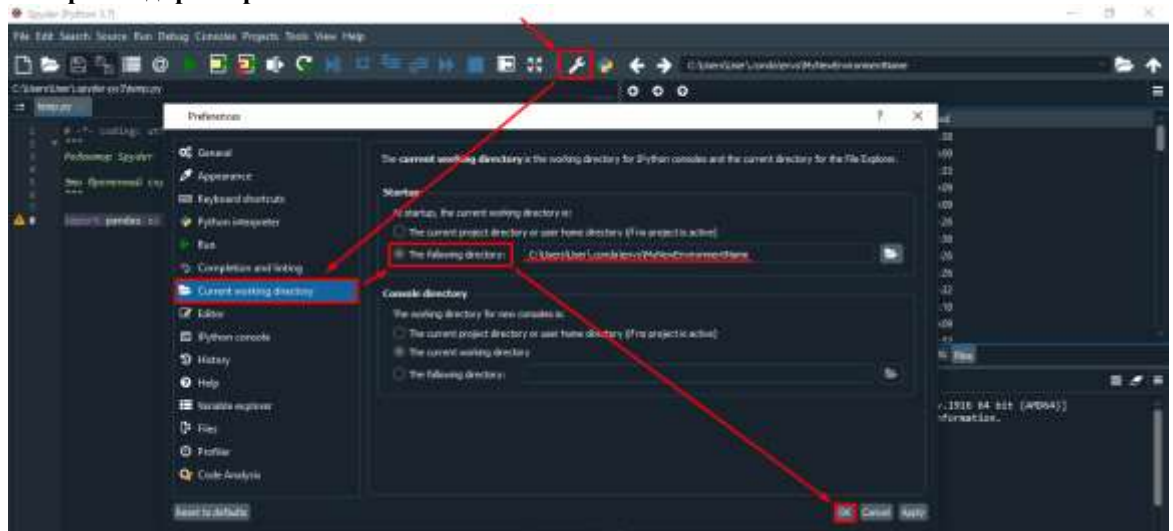
Настройка среды для Spyder

#### 1. Настройка интерпретатора



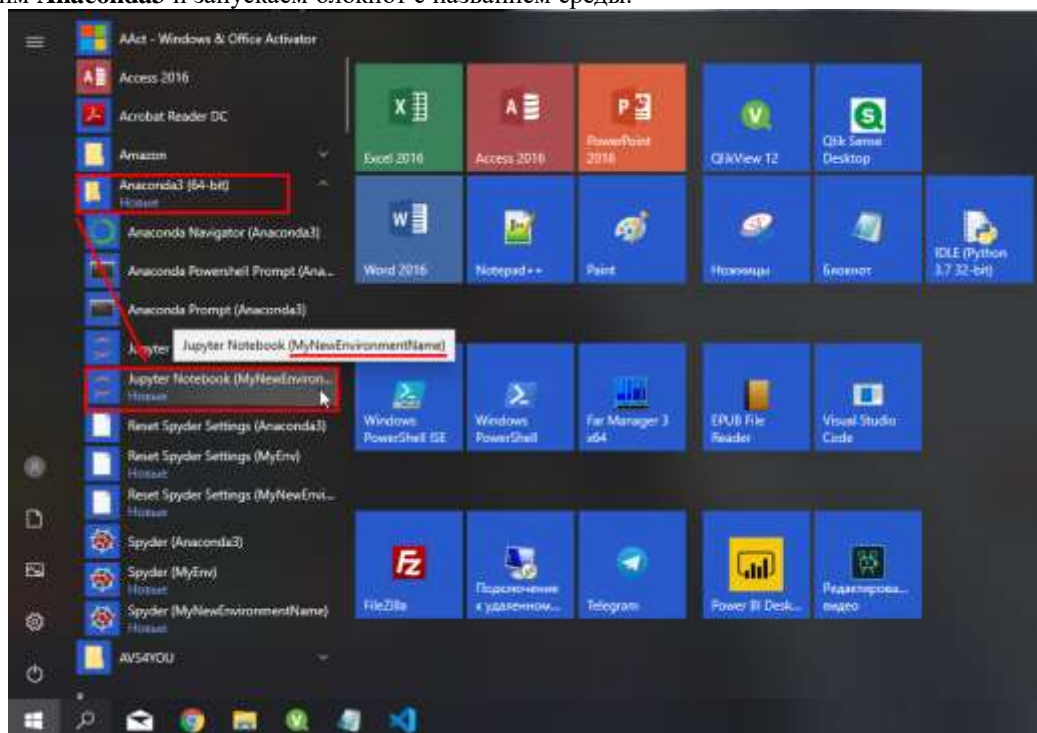


### Настройка директории



Как открыть Jupyter Notebook в новой среде MyNewEnvironmentName

Для того, чтобы запустить **Jupyter Notebook** в созданной среде **MyNewEnvironmentName**, в поиске находим **Anaconda3** и запускаем блокнот с названием среды:

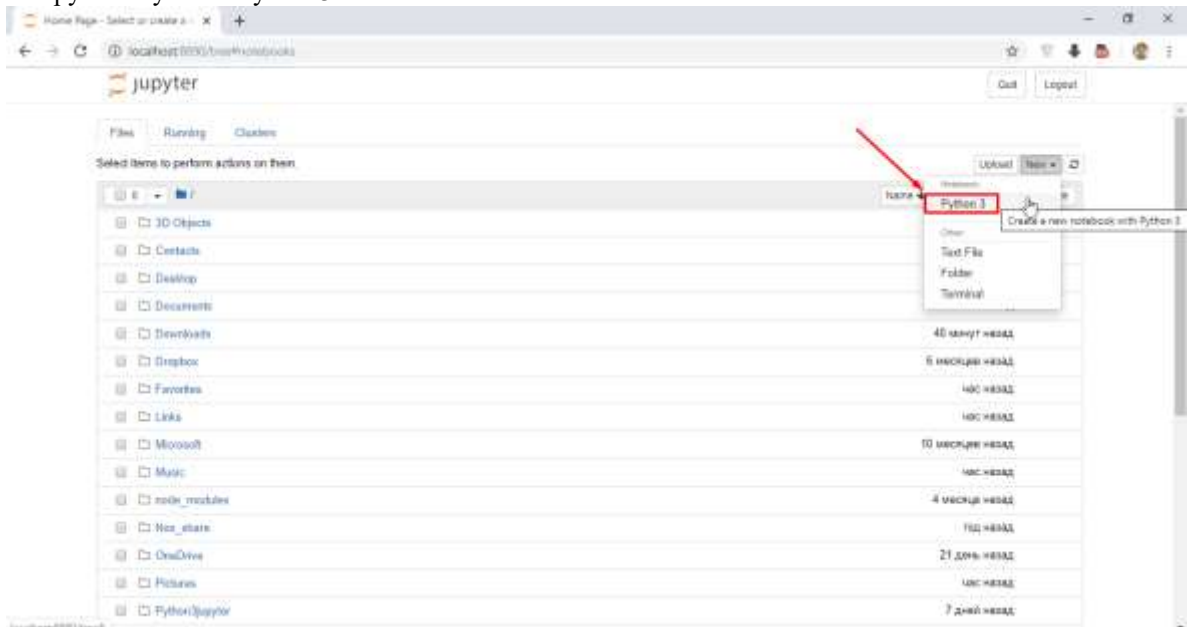


Появится консольное окошко — это движок Jupyter Notebook, который работает в фоновом режиме:

```
Jupyter Notebook (MyNewEnvironmentName)
[I 23:57:47.710 NotebookApp] The port 8888 is already in use, trying another port.
[I 23:57:47.710 NotebookApp] The port 8889 is already in use, trying another port.
[I 23:57:47.757 NotebookApp] Serving notebooks from local directory: C:\Users\User
[I 23:57:47.757 NotebookApp] The Jupyter Notebook is running at:
[I 23:57:47.757 NotebookApp] http://localhost:8890/?token=57cb326d88828682f9384226df11d58fff936a7547108cc6
[I 23:57:47.757 NotebookApp] or http://127.0.0.1:8890/?token=57cb326d88828682f9384226df11d58fff936a7547108cc6
[I 23:57:47.757 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 23:57:47.969 NotebookApp]

To access the notebook, open this file in a browser:
File:///C:/Users/User/AppData/Roaming/jupyter/runtime/nbserver-9888-open.html
Or copy and paste one of these URLs:
http://localhost:8890/?token=57cb326d88828682f9384226df11d58fff936a7547108cc6
or http://127.0.0.1:8890/?token=57cb326d88828682f9384226df11d58fff936a7547108cc6
```

В Jupyter запускаем Python 3:



Для того, чтобы убедиться в какой среде мы работаем, можно вбить ряд команд (ниже приведен текст этих команд для Python 3):

```
jupyter MyNewEnvironmentName_Python_Example Last Checkpoint: несколько секунд назад (unsaved changes)
Python 3

In [15]: import sys
         print(sys.version)
         print(sys.base_prefix)

3.7.5 (default, Oct 31 2019, 15:18:51) [MSC v.1010 64 bit (AMD64)]
C:\Users\User\.conda\envs\MyNewEnvironmentName

In [16]: print('\n'.join(sys.modules.keys()))

sys
builtins
_frozen_importlib
_imp
_impthread
_warnings
_weakref
zipimport
_frozen_importlib_external
_sio
marshal
nt
winreg
encodings
codecs
codecs
encodings.aliases
encodings.utf_8
_signal
```

Узнать среду, в которой работает Jupyter Notebook:

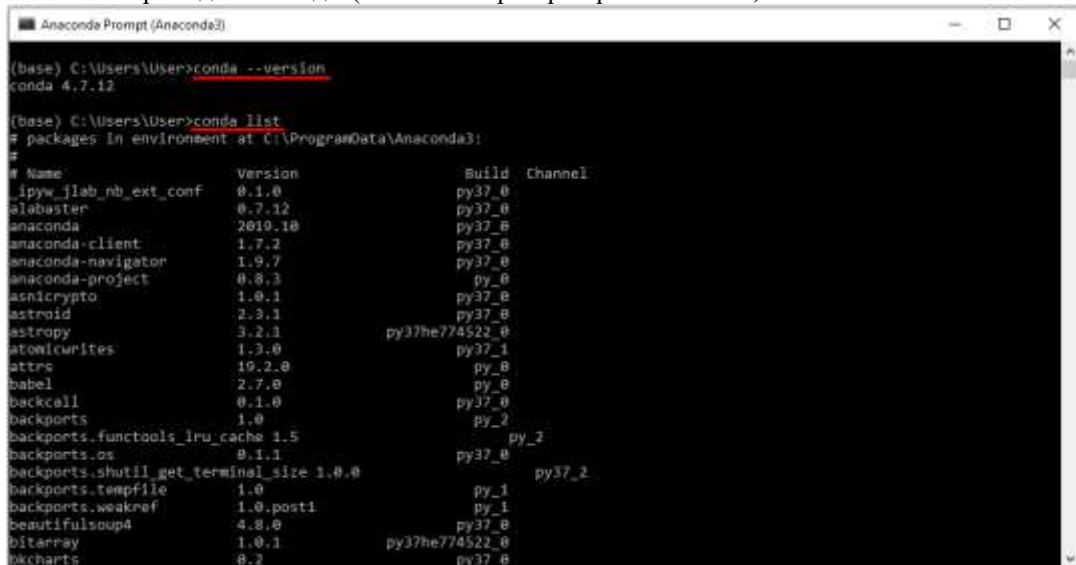
```
import sys
print(sys.version)
print(sys.base_prefix)
Получить список модулей, доступных в Env:
print('\n'.join(sys.modules.keys()))
```

Anaconda3 Prompt cmd Conda Command — Запуск команд через консоль

Для того, чтобы использовать команды conda через командную строку (cmd), необходимо запустить программу **Anaconda Prompt (Anaconda3)**

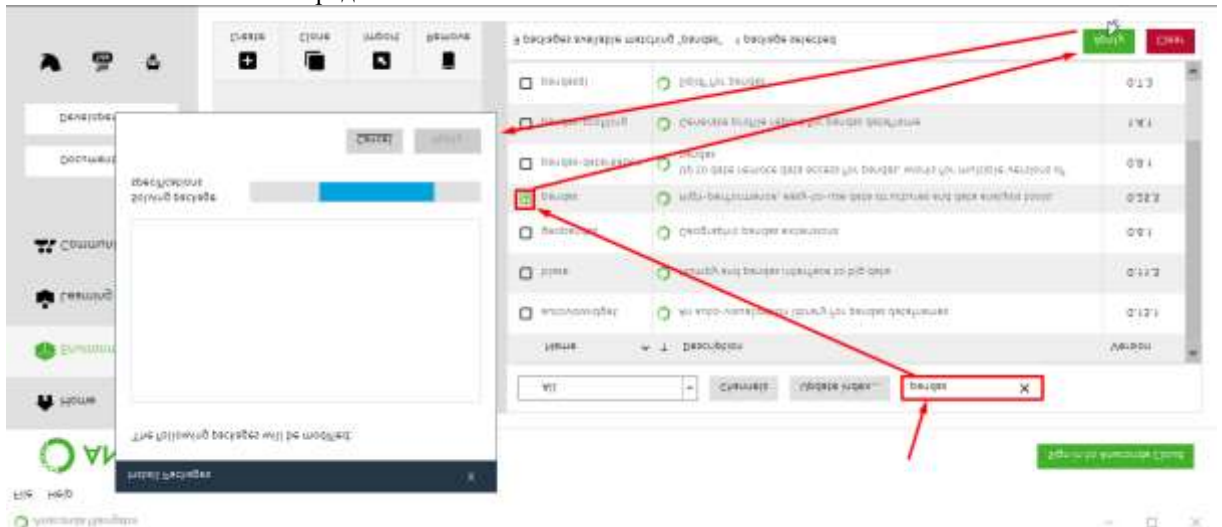


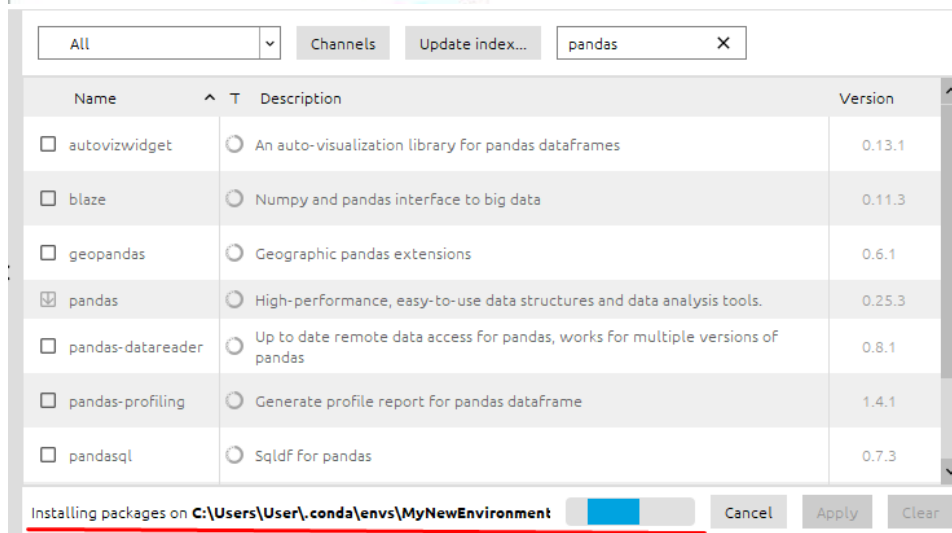
Можете набрать две команды (в качестве проверки работы conda):



Установка новой библиотеки (пакета) в среду

Пакеты управляются отдельно для каждой среды. Изменения, которые вы вносите в пакеты, применяются только к активной среде.





Исполняемые файлы в среде Conda

**python.exe** — исполняемый файл Python для приложений командной строки. Так, например, если вы находитесь в каталоге Example App, вы можете выполнить его: `python.exe exampleapp.py`

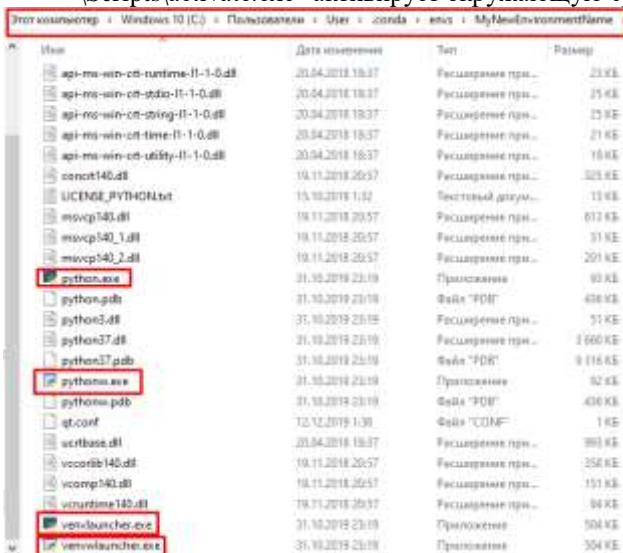
**pythonw.exe** — исполняемый файл Python для приложений с графическим интерфейсом или приложений без интерфейса пользователя

**venvlauncher.exe** —

**venvwlauncher.exe** —

`\Scripts` — исполняемые файлы, являющиеся частью установленных пакетов. После активации среды этот каталог добавляется в системный путь, поэтому исполняемые файлы становятся доступными без их полного пути.

`\Scripts\activate.exe` - активирует окружающую среду



Источники:

- <https://protostar.space/why-you-need-python-environments-and-how-to-manage-them-with-conda>
- [https://kapeli.com/cheat\\_sheets/Conda.docset/Contents/Resources/Documents/index](https://kapeli.com/cheat_sheets/Conda.docset/Contents/Resources/Documents/index)