

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Мурманский арктический государственный университет»
(ФГБОУ ВО «МАГУ»)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.03.02 Python для извлечения и обработки данных

(название дисциплины (модуля) в соответствии с учебным планом)

**основной профессиональной образовательной программы
по направлению подготовки**

01.03.02 Прикладная математика и информатика

(код и наименование направления подготовки)

Управление данными и машинное обучение

(наименование направленности (профиля / профилей) / магистерской программы))

высшее образование – бакалавриат

уровень профессионального образования: высшее образование – бакалавриат / высшее образование – специалитет,
магистратура / высшее образование – подготовка кадров высшей квалификации

бакалавр

квалификация

очная

форма обучения

2021

год набора

Составитель(и):

Ляш Олег Иванович,
доцент, канд. пед. наук, доцент
кафедры математики, физики
и информационных технологий

Утверждено на заседании кафедры
математики, физики и информационных
технологий факультета
математических и естественных наук
(протокол № 07 от 12.04.2021)

Переутверждено на заседании кафедры
математики, физики и информационных
технологий факультета
математических и естественных наук
(протокол № 09 от 02.07.2021)

Зав. кафедрой _____ Лазарева И.М.
подпись Ф.И.О.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ) — формирование у студентов совокупности знаний и представлений о технологиях сбора и анализа данных, получаемых с различных сайтов. В ходе прохождения дисциплины студенты знакомятся с основами анализа данных, актуализируют и расширяют свои знания о средствах сбора информации.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

В результате освоения дисциплины (модуля) формируются следующие компетенции:

ПК-2 способен работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности

ПК-3 Способен к разработке и применению алгоритмических и программных решений в области системного и прикладного программного обеспечения

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-2 способен работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности	ПК-2.1 Формулирует задачи в рамках проекта и определяет ожидаемые результаты ПК-2.2 Обеспечивает модульность выполнения задачи с учетом имеющихся ресурсов ПК-2.3 Обеспечивает пользовательскую привлекательность создаваемого программного продукта	Знать: – основные подходы к анализу данных; – технологии сбора данных; – технологии визуализации данных. Уметь: – эффективно использовать программные средства для анализа данных; – использовать отчёты систем, генерирующих потоки данных.
ПК-3 Способен к разработке и применению алгоритмических и программных решений в области системного и прикладного программного обеспечения	ПК-3.1 Разрабатывает алгоритм решения поставленной задачи выбранным методом ПК-3.2 Выбирает и обосновывает выбор языковой среды ПК-3.3 Использует современную языковую среду для реализации сложных алгоритмов ПК-3.4 Решает задачу тестирования программного продукта	Владеть: – навыками обработки данных; – навыками сбора данных; – навыками визуализации данных.

3. УКАЗАНИЕ МЕСТА ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОСНОВНОЙ ПРОФЕССИОНАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина (модуль) «Python для извлечения и обработки данных» относится к части, формируемой участниками образовательных отношений, образовательной программы по направлению подготовки 01.03.02 Прикладная математика и информатика, направленность (профиль) Управление данными и машинное обучение.

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Общая трудоемкость дисциплины (модуля) составляет 3 зачетные единицы или 108 часов, из расчета 1 ЗЕ = 36 часов.

Курс	Семестр	Трудоемкость в ЗЕ	Общая трудоемкость (час)	Контактная работа			Всего контактных часов	Из них:		Кол-во часов на СРС		Кол-во часов на контроль	Форма контроля
				ЛК	ПР	ЛБ		В интерактивной форме	В форме практической подготовки	Общее количество часов на СРС	Из них – на курсовую работу		
3	6	3	108	14	-	22	36	8	12	72	-		Зачёт

В интерактивных формах часы используются в виде обсуждения вопросов по теме дисциплины на лекционных занятиях.

Практическая подготовка реализуется в виде решения практических задач.

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ), СТРУКТУРИРОВАННОЕ ПО ТЕМАМ (РАЗДЕЛАМ) С УКАЗАНИЕМ ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ

№ п/п	Наименование раздела, темы	Контактная работа			Всего контактных часов	Из них:		Кол-во часов на СРС	Кол-во часов на контроль
		ЛК	ПР	ЛБ		В интерактивной форме	В форме практической подготовки		
1	Введение в анализ данных	4		6	10				
2	Источники данных	4		8	12	4	6	32	
3	Инструменты обработки данных	6		8	14	4	6	40	
	ИТОГО	14		22	36	8	12	72	

Содержание дисциплины (модуля)

Тема 1. Введение в анализ данных

Деревья решений. Самоорганизующиеся карты признаков. Ассоциативные правила. Факторный анализ. Линейная регрессия. Логистическая регрессия. Кластеризация данных: основы кластерного анализа; иерархический кластерный анализ; двумерная визуализация; трехмерная визуализация.

Тема 2. Источники данных

Генерация отчетов web-сервером. Системы сбора данных Яндекс. Системы сбора данных Google. Выгрузка данных Discord. Выгрузка данных Telegram. Выгрузка данных MS teams. Выгрузка данных Viber. Выгрузка данных whatsapp.

Тема 3. Инструменты обработки данных

Интерфейс Scilab. Типы данных. Действия над матрицами: создание матриц; генерация матриц специального вида; доступ к ячейкам; произведение матриц. Обработка данных средствами табличного процессора. Использование Google Data Studio. Использование Yandex Data Lens.

6. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ, НЕОБХОДИМОГО ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ).

Основная литература:

1. Низаметдинов, Ш. У. Анализ данных : учебное пособие / Ш. У. Низаметдинов, В. П. Румянцев. — Москва : НИЯУ МИФИ, 2012. — 288 с. — ISBN 978-5-7262-1687-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/75847> — Режим доступа: для авториз. Пользователей.
2. Алексеев, Д. С. Технологии интеллектуального анализа данных : учебное пособие / Д. С. Алексеев. — Кострома : КГУ им. Н.А. Некрасова, 2020. — 141 с. — ISBN 978-5-8285-1083-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/160082> — Режим доступа: для авториз. Пользователей.
3. Крутиков, В.Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Кемеровский государственный университет. — Кемерово : Кемеровский государственный университет, 2014. — 138 с. : ил. — Режим доступа: по подписке. — URL: <https://biblioclub.ru/index.php?page=book&id=278426> — Библиогр. в кн. — ISBN 978-5-8353-1770-7. — Текст : электронный.

Дополнительная литература:

4. Жуковский, О.И. Информационные технологии и анализ данных : учебное пособие / О.И. Жуковский ; Томский Государственный университет систем управления и радиоэлектроники (ТУСУР). — Томск : Эль Контент, 2014. — 130 с. : схем., ил. — Режим доступа: по подписке. — URL:

<https://biblioclub.ru/index.php?page=book&id=480500> – Библиогр.: с. 126. – ISBN 978-5-4332-0158-3. – Текст : электронный.

5. Агалаков, С.А. Статистические методы анализа данных : учебное пособие : [16+] / С.А. Агалаков ; Омский государственный университет им. Ф. М. Достоевского. – Омск : Омский государственный университет им. Ф.М. Достоевского, 2017. – 92 с. : табл., граф., схем., ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=562918> – Библиогр. в кн. – ISBN 978-5-7779-2187-1. – Текст : электронный.
6. Орешков, В. И. Интеллектуальный анализ данных : учебное пособие / В. И. Орешков. — Рязань : РГРТУ, 2017. — 160 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/168028> — Режим доступа: для авториз. пользователей.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

В образовательном процессе используются:

- учебные аудитории для проведения учебных занятий, оснащенные оборудованием и техническими средствами обучения: учебная мебель, ПК, оборудование для демонстрации презентаций, наглядные пособия;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду МАГУ.

7.1 ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО И СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ:

7.1.1. Лицензионное программное обеспечение отечественного производства.

7.1.2. Лицензионное программное обеспечение зарубежного производства:

- Операционная система: MS Windows версии 7 и выше
- Программные средства, входящие в состав офисного пакета: MS Office (Word, Excel, Access, Publisher, PowerPoint)
- Программы для просмотра документов: Adobe Acrobat Reader

7.1.3. Свободно распространяемое программное обеспечение отечественного производства:

7.1.4. Свободно распространяемое программное обеспечение зарубежного производства:

- Программные средства, входящие в состав офисного пакета: LibreOffice (Writer, Calc, Base, Impress, Draw)
- Текстовые редакторы: Notepad ++
- Графические редакторы: InkScape, Gimp
- Системы программирования: Pascal ABC, Python IDLE
- Браузеры: Mozilla Firefox

7.2 ЭЛЕКТРОННО-БИБЛИОТЕЧНЫЕ СИСТЕМЫ:

- ЭБС издательства «Лань» <https://e.lanbook.com/>
- ЭБС издательства «Юрайт» <https://biblio-online.ru/>
- ЭБС «Университетская библиотека онлайн» <https://biblioclub.ru/>

7.3 СОВРЕМЕННЫЕ ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ

- Информационно-аналитическая система SCIENCE INDEX
- Электронная база данных Scopus
- Базы данных компании CLARIVATE ANALYTICS

7.4. ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ

- Справочно-правовая информационная система Консультант Плюс <http://www.consultant.ru/>
- ООО «Современные медиа технологии в образовании и культуре» <http://www.informio.ru/>

8. ИНЫЕ СВЕДЕНИЯ И МАТЕРИАЛЫ НА УСМОТРЕНИЕ ВЕДУЩЕЙ КАФЕДРЫ

Не предусмотрено

9. ОБЕСПЕЧЕНИЕ ОБРАЗОВАНИЯ ДЛЯ ЛИЦ С ОВЗ

Для обеспечения образования инвалидов и лиц с ограниченными возможностями здоровья реализация дисциплины может осуществляться в адаптированном виде, с учетом специфики освоения и дидактических требований, исходя из индивидуальных возможностей и по личному заявлению обучающегося.