

Компонент ОПОП 09.04.01 Информатика и вычислительная техника
направленность (профиль) Компьютерный анализ и интерпретация данных.

Data Science
наименование ОПОП

Б1.О.14
шифр дисциплины

РАБОЧАЯ ПРОГРАММА

Дисциплины
(модуля)

Задачи обработки, анализа и интерпретации данных

Разработчик:

Золотов О.В.
ФИО

доцент
должность

канд. физ.-мат. наук
ученая степень,
звание

Утверждено на заседании кафедры
Цифровых технологий, математики
и экономики
наименование кафедры

протокол № 8 от 23.05.2023

И.о. заведующего кафедрой
ЦТМиЭ

_____ Мотина Т.Н.
подпись _____ ФИО

Мурманск
2023

Пояснительная записка

Объем дисциплины 7 з.е.

1. Результаты обучения по дисциплине (модулю), соотнесенные с индикаторами достижения компетенций, установленными образовательной программой

Компетенции	Индикаторы достижения компетенций	Результаты обучения по дисциплине (модулю)
<p>ОПК-2 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач</p>	<p>ИД-1опк-2 Понимает принципы и применяет методы разработки алгоритмов и программных средств</p> <p>ИД-2опк-2 Использует современные интеллектуальные технологии при разработке алгоритмов и программных средств</p> <p>ИД-3опк-2 Решает профессиональные задачи путем разработки оригинальных алгоритмов и программных средств</p>	<p>Знать: принципы и методы разработки алгоритмов и программных средств, современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач в области обработки, анализа и интерпретации данных;</p> <p>общие принципы осуществления поиска, сравнения и оценки исследований, методы их проведения и проверки применимости для решения задач в области обработки, анализа и интерпретации данных;</p> <p>общие принципы работы программно-аппаратных комплексов обработки информации и автоматизированного проектирования, анализа подходов к разработке их компонент, инструментальные средства разработки, применяемые при решении задач в области обработки, анализа и интерпретации данных;</p> <p>функциональные требования к прикладному программному обеспечению для решения актуальных задач обработки, анализа и интерпретации данных предприятий отрасли, национальные стандарты обработки информации и автоматизированного проектирования.</p>
<p>ОПК-4 Способен применять на практике новые научные принципы и методы исследований</p>	<p>ИД-1опк-4 Оценивает новые научные принципы и их применимость для решения практических задач</p> <p>ИД-2опк-4 Осуществляет поиск, сравнение, оценку методов исследований</p> <p>ИД-3опк-4 Применяет новые методы проведения исследований для решения практических задач профессиональной деятельности</p>	<p>Уметь: осуществлять обоснованный выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач в области обработки, анализа и интерпретации данных;</p> <p>формулировать принципы, осуществлять поиск, сравнение, оценку методов исследований, оценивать их применимость для решения практических задач в области обработки, анализа и интерпретации данных;</p>
<p>ОПК-6 Способен разрабатывать компоненты программно-аппаратных комплексов</p>	<p>ИД-1опк-6 Понимает принципы работы программно-аппаратных комплексов обработки</p>	<p>выполнять анализ подходов к разработке компонент программно-аппаратных комплексов обработки информации и автоматизированного проектирования с учетом принципов их работы, использовать инструментальные средства разработки при решении задач в области</p>

<p>обработки информации и автоматизированного проектирования</p>	<p>информации и автоматизированного проектирования ИД-2опк-6 Анализирует существующие подходы к разработке компонент программно-аппаратных комплексов обработки информации и автоматизированного проектирования ИД-3опк-6 Использует инструментальные средства разработки компонент программно-аппаратных комплексов</p>	<p>обработки, анализа и интерпретации данных; приводить с учетом специфики зарубежные комплексы обработки информации, анализа и интерпретации данных в соответствие с национальными требованиями и потребностями отечественных предприятий, интегрировать с отраслевыми информационными системами. Владеть: методами разработки алгоритмов и программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач в области обработки, анализа и интерпретации данных;</p> <p>приемами поиска, сравнения и оценки методов исследования, их проведения и оценивания на предмет применимости для решения практических задач профессиональной деятельности в области обработки, анализа и интерпретации данных;</p> <p>навыком применения приемов анализа подходов к разработке компонент программно-аппаратных комплексов обработки информации и автоматизированного проектирования с учетом принципов их работы, использования инструментальных средств разработки при решении задач в области обработки, анализа и интерпретации данных;</p>
<p>ОПК-7 Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий</p>	<p>ИД-1опк-7 Понимает специфику зарубежных комплексов обработки информации и автоматизированного проектирования ИД-2опк-7 Анализирует цифровые потребности отечественных предприятий ИД-3опк-7 Адаптирует комплексы обработки информации и автоматизированного проектирования к нуждам конкретного предприятия</p>	<p>методами анализа и настройки интерфейса, разработки пользовательских шаблонов, подключения библиотек, добавления новых функций при решении задач в области обработки, анализа и интерпретации данных.</p>

2. Содержание дисциплины (модуля)

курс 1, семестр 2

Тема 1. Предмет курса. Данные, информация, знания. Задачи обработки, анализа и интерпретации данных, история развития дисциплины, основные направления исследований и разработок (2 ак. ч.)

Тема 2. Выполнение многомерных вычислений (на примере библиотеки научных вычислений NumPy) (6 ак. ч.)

Тема 3. Компьютерный анализ и обработка данных (на примере библиотек Pandas и хагау) (6 ак. ч.)

Тема 4. Визуализация данных (на примере библиотек matplotlib и seaborn) (4 ак. ч.)

курс 2, семестр 3

Тема 5. Обработка данных (на примере возможностей библиотеки машинного обучения Scikit-Learn) (6 ак. ч.)

Тема 6. Организация конвейеров обработки данных на стеке ELK: компонент Elasticsearch (2 ак. ч.)

Тема 7. Организация конвейеров обработки данных на стеке ELK: компонент Logstash (2 ак. ч.)

Тема 8. Организация конвейеров обработки данных на стеке ELK: компоненты Beats и Kibana (2 ак. ч.)

Тема 9. Организация параллельной обработки данных (на примере *Apache Spark* и PySpark) (6 ак. ч.)

3. Перечень учебно-методического обеспечения дисциплины (модуля)

- мультимедийные презентационные материалы по дисциплине (модулю) представлены в электронном курсе в ЭИОС МАУ;
- методические указания к выполнению лабораторных работ представлены в электронном курсе в ЭИОС МАУ;
- методические материалы для обучающихся по освоению дисциплины (модуля) представлены на официальном сайте МАУ в разделе «Информация по образовательным программам, в том числе адаптированным».

4. Фонд оценочных средств по дисциплине (модулю)

Является отдельным компонентом образовательной программы, разработан в форме отдельного документа, включает в себя:

- перечень компетенций с указанием этапов их формирования в процессе освоения дисциплины (модуля);
- задания текущего контроля;
- задания промежуточной аттестации;
- задания внутренней оценки качества образования.

5. Перечень основной и дополнительной учебной литературы (печатные издания, электронные учебные издания и (или) ресурсы электронно-библиотечных систем)

Основная литература:

1. Лейхтер, С. В., Анализ данных / Лейхтер, С. В., С. Н. Чуканов, И. С. Чуканов, И. В. Широков. – Омск : Омский государственный университет им. Ф.М. Достоевского, 2022. – 108 с. – Текст : электронный. – URL: <https://elibrary.ru/item.asp?id=49448527>
2. Чернавин, П.Ф., Машинное обучение на основе задач математического программирования / П. Ф. Чернавин, Д. Н. Гайнанов, В. Н. Панкращенко, Ф. П. Чернавин, Н. П. Чернавин. – Москва : Федеральное государственное унитарное предприятие "Академический научно-издательский, производственно-полиграфический и книгораспространительский центр "Наука", 2021. – 128 с. – Текст : электронный. – URL: <https://elibrary.ru/item.asp?id=47402017>
3. Гусев, С. И., Цифровая обработка данных радиолокационного зондирования Земли из космоса / С. И. Гусев, Н. А. Егоскин, В. В. Еремеев, А. Е. Кузнецов, А. Э. Москвитин, В. А. Ушенкин. – Москва : Общество с ограниченной ответственностью Издательство "КУРС", 2021. – 248 с. – Текст : электронный. – URL: <https://elibrary.ru/item.asp?id=47697294>

Дополнительная литература:

4. Евдокимов, Ю. К., Автоматизированный сбор и цифровая обработка данных в измерительных системах / Ю. К. Евдокимов, Е. С. Денисов, Д. В. Шахтурин, О. В. Шиндор. – Казань : Казанский национальный исследовательский технологический университет, 2012. – 163 с. – Текст : электронный. – URL: <https://elibrary.ru/item.asp?id=47402017>
5. NumPy User Guide. – Text : electronic // NumPy documentation : [site]. – [S. I.], 2022. – URL: <https://numpy.org/doc/stable/numpy-user.pdf>
6. NumPy Tutorial. – Text : electronic // w3school : [site]. – [S. I., 2022]. – URL: https://www.w3schools.com/python/numpy/numpy_intro.asp
7. User Guide. – Text : electronic // Pandas documentation : [site]. – [S. I.], 2022. – URL: https://pandas.pydata.org/docs/user_guide/index.html
8. User Guide. – Text : electronic // Xarray documentation : [site]. – [S. I.], 2022. – URL: <https://docs.xarray.dev/en/stable/user-guide/index.html>
9. User Guide. – Text : electronic // Matplotlib 3.6.0 documentation : [site]. – [S. I.], 2022. – URL: <https://matplotlib.org/stable/users/index.html>
10. Matplotlib Tutorial. – Text : electronic // w3school : [site]. – [S. I., 2022]. – URL: https://www.w3schools.com/python/matplotlib_intro.asp
11. Tutorials. – Text : electronic // PyGMT : [site]. – [S. I., 2022]. – URL: <https://www.pygmt.org/latest/tutorials/index.html>
12. User Guide. – Text : electronic // scikit-learn : [site]. – [S. I.], 2022. – URL: https://scikit-learn.org/stable/user_guide.html
13. Scikit Learn Tutorial. – Text : electronic // TutorialsPoint : [site]. – [S. I.], 2022. – URL: https://www.tutorialspoint.com/scikit_learn/scikit_learn_tutorial.pdf
14. PySpark Documentation. – Text : electronic // Apache Spark - Unified engine for large-scale data analytics : [site]. – [S. I.], 2022. – URL: <https://spark.apache.org/docs/latest/api/python/>
15. ELK Stack Tutorial: What is Kibana, Logstash & Elasticsearch?. – Text : electronic // Guru99 [site]. – [S. I.], 2022. – URL: <https://www.guru99.com/elk-stack-tutorial.html>
16. Horovits, D., The Complete Guide to the ELK Stack / D. Horovits. – Text : electronic // Logz.io: Cloud Observability for Engineer : [site]. – [S. I.], 2022. – URL: <https://logz.io/learn/complete-guide-elk-stack/>
17. Logstash Tutorial. – Text : electronic // TutorialsPoint : [site]. – [S. I.], 2022. – URL: <https://www.tutorialspoint.com/logstash/index.htm>

6. Профессиональные базы данных и информационные справочные системы

1. eLIBRARY.RU : научная электронная библиотека. – Москва, 2000- . – URL: <https://www.elibrary.ru/> . – Режим доступа: научная электронная библиотека eLIBRARY.RU, после регистрации.
2. Kaggle : AI & ML community : site. – 2024. – URL: <https://www.kaggle.com/>
3. КонсультантПлюс : справочная правовая система : сайт. – Москва, 1997- . – URL: <https://www.consultant.ru>

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства

1. свободное программное обеспечение с открытым исходным кодом интерпретатор Python3;
2. свободное программное обеспечение с открытым исходным кодом pip – стандартный установщик и менеджер зависимостей Python-пакетов;

3. браузер, поставляемый вместе с операционной системой (либо свободное программное обеспечение Mozilla FireFox или Google Chrome);
4. виртуальная машина для запуска гостевых операционных систем (VirtualBox или Hyper-V, по доступности);
5. свободное программное обеспечение - операционная система Linux (на базе дистрибутива Rocky Linux).
6. свободное программное обеспечение – Elasticsearch, Logstash, Kibana
7. свободное программное обеспечение – Apache Spark
8. свободное программное обеспечение – Python-библиотеки NumPy, Pandas, xarray, matplotlib, seaborn, Scikit-Learn.

8. Обеспечение освоения дисциплины лиц с инвалидностью и ОВЗ

Обучающиеся из числа инвалидов и лиц с ОВЗ обеспечиваются печатными и (или) электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

9. Материально-техническое обеспечение дисциплины (модуля) представлено в приложении к ОПОП «Материально-технические условия реализации образовательной программы» и включает:

- учебные аудитории для проведения учебных занятий, предусмотренных программой магистратуры, оснащенные оборудованием и техническими средствами обучения;

- помещения для самостоятельной работы обучающихся, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде МАУ.

Допускается замена оборудования его виртуальными аналогами.

10. Распределение трудоемкости по видам учебной деятельности

Таблица 1 - Распределение трудоемкости

Вид учебной деятельности	Распределение трудоемкости дисциплины (модуля) по формам обучения											
	Очная				Очно-заочная				Заочная			
	Семестр			Всего часов	Семестр			Всего часов	Семестр/Курс			Всего часов
	2	3	-		-	-	-		-	-	-	
Лекции	18	18	-	36	-	-	-	-	-	-	-	-
Практические занятия	-	-	-	-	-	-	-	-	-	-	-	-
Лабораторные работы	24	24	-	48	-	-	-	-	-	-	-	-
Самостоятельная работа	66	66	-	132	-	-	-	-	-	-	-	-
Подготовка к промежуточной аттестации ¹	-	36	-	36	-	-	-	-	-	-	-	-
Всего часов по дисциплине	108	144	-	252	-	-	-	-	-	-	-	-
/ из них в форме практической подготовки ²	24	24	-	48	-	-	-	-	-	-	-	-

Формы промежуточной аттестации и текущего контроля

Экзамен	-	+	-	1	-	-	-	-	-	-	-	-
Зачет/зачет оценкой	+/-	-	-	1/-	-	-	-	-	-	-	-	-
Курсовая работа (проект)	-(-)	-	-	-	-	-	-	-	-	-	-	-
Количество расчетно-графических работ	1	1		2	-	-	-	-	-	-	-	-
Количество контрольных работ	-		-	-	-	-	-	-	-	-	-	-
Количество рефератов	-	-	-	-	-	-	-	-	-	-	-	-
Количество эссе	-	-	-	-	-	-	-	-	-	-	-	-

¹ Для экзамена очной и очно-заочной формы обучения - 36 часов, для экзамена заочной формы обучения - 9 часов, для зачета заочной формы обучения - 4 часа.

² Организуется при реализации учебных дисциплин (модулей) путем проведения практических занятий, лабораторных работ и иных аналогичных видов учебной деятельности, предусматривающих участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Практическая подготовка может включать в себя отдельные занятия лекционного типа, которые предусматривают передачу учебной информации обучающимся, необходимой для последующего выполнения работ, связанных с будущей профессиональной деятельностью.

Перечень лабораторных работ по формам обучения

№ п/п	Темы лабораторных работ
1	2
	Очная форма
	семестр 2
1	Л/р №1.1. Библиотека NumPy в задачах анализа, обработки и интерпретации данных
2	Л/р №1.2. Библиотека NumPy в задачах анализа, обработки и интерпретации данных
3	Л/р №1.3. Библиотека NumPy в задачах анализа, обработки и интерпретации данных
4	Л/р №1.4. Библиотека NumPy в задачах анализа, обработки и интерпретации данных
5	Л/р №2.1. Библиотека Pandas в задачах анализа, обработки и интерпретации данных
6	Л/р №2.2. Библиотека Pandas в задачах анализа, обработки и интерпретации данных
7	Л/р №2.3. Библиотека Pandas в задачах анализа, обработки и интерпретации данных
8	Л/р №2.4. Библиотека Pandas в задачах анализа, обработки и интерпретации данных
9	Л/р №3. Библиотека хarray в задачах анализа, обработки и интерпретации данных
10	Л/р №4.1. Визуализация данных с использованием библиотек matplotlib и seaborn в задачах анализа, обработки и интерпретации данных
11	Л/р №4.2. Визуализация данных с использованием библиотек matplotlib и seaborn в задачах анализа, обработки и интерпретации данных
12	Л/р №4.3. Визуализация данных с использованием библиотек matplotlib и seaborn в задачах анализа, обработки и интерпретации данных
	семестр 3
1	Л/р №1.1. Библиотека машинного обучения Scikit-Learn
2	Л/р №1.2. Библиотека машинного обучения Scikit-Learn
3	Л/р №1.3. Библиотека машинного обучения Scikit-Learn
4	Л/р №2.1. Стек ELK – Elasticsearch
5	Л/р №2.2. Стек ELK – Elasticsearch
6	Л/р №3.1. Стек ELK – Logstash
7	Л/р №3.2. Стек ELK – Logstash
8	Л/р №4.1. Стек ELK – Kibana
9	Л/р №4.2. Стек ELK – Kibana
10	Л/р №5.1. Apache Spark и PySpark
11	Л/р №5.2. Apache Spark и PySpark
12	Л/р №5.3. Apache Spark и PySpark
	Заочная форма
-	Заочная форма обучения не реализуется

Перечень практических занятий по формам обучения

№ п/п	Темы практических занятий
1	2
	Очная форма
-	Практические занятия не предусмотрены учебным планом
	Заочная форма
-	Заочная форма обучения не реализуется

Перечень примерных тем курсовой работы /курсового проекта

№ п/п	Темы курсовой работы /проекта
1	2
-	Курсовая работа / проект не предусмотрены учебным планом