

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

Московский государственный университет имени М.В. Ломоносова  
Факультет Вычислительной математики и кибернетики



УТВЕРЖДАЮ  
Декан факультета ВМК  
/ Соколов И.А. /  
2023 г.

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Наименование дисциплины (модуля):

Межфакультетский курс «Основы машинного обучения»

код и наименование дисциплины (модуля)

Уровень высшего образования:

бакалавриат, магистратура, специалитет

указывается: бакалавриат, магистратура или специалитет

Направление подготовки (специальность):

(код и название направления/специальности)

Направленность (профиль) ОПОП:

(если дисциплина (модуль) относится к вариативной части программы)

Форма обучения:

Очная с использованием электронных образовательных технологий

очная, очно-заочная, заочная

Рабочая программа рассмотрена и одобрена  
на заседании Ученого совета факультета  
(протокол № 1, 20.01.2023)

Москва 2023

***На обратной стороне титула:***

Рабочая программа дисциплины (модуля) разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки программ бакалавриата, магистратуры, специалитета.

1. Место дисциплины (модуля) в структуре ОПОП реализуется в рамках МФК.
2. Входные требования для освоения дисциплины (модуля), предварительные условия: не требуются.
3. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников.

Компетенции выпускников (коды)	Индикаторы (показатели) достижения компетенций	Планируемые результаты обучения по дисциплине (модулю), сопряженные с компетенциями
<p><b>Новая УК ОС МГУ</b> Способен осуществлять коммуникацию, поиск, обработку и анализ данных с применением цифровых инструментов, в том числе с элементами программирования и технологий искусственного интеллекта.</p>	<p>УК-N (Ин.1ук) Знает основные понятия и методы программирования, виды и форматы хранения данных, этапы проведения анализа данных и основные понятия искусственного интеллекта и машинного обучения</p>	<p><b>Знать:</b> Основные понятия программирования Этапы решения задач и применением средств вычислительной техники Виды и форматы хранения данных Этапы проведения анализа данных Стандартные алгоритмы обработки и анализа данных Основные понятия искусственного интеллекта и машинного обучения</p> <p><b>Уметь</b> Применять средства и технологии программирования для решения задач связанных с анализом данных Проводить поиск, обработку и анализ данных для решения профессиональных задач, в том числе открытых данных Использовать технологии искусственного интеллекта и машинного обучения для решения профессиональных задач</p> <p><b>Владеть:</b> методами и технологиями использования средств вычислительной техники для решения профессиональных задач, в том числе с применением технологий искусственного интеллекта</p>
	<p>УК-N (Ин.2ук) Умеет применять стандартные алгоритмы и средства программирования для решения задач, связанных с анализом данных, в том числе с использованием технологий искусственного интеллекта и машинного обучения</p>	
	<p>УК-N (Ин.3ук) Владеет методами и технологиями использования средств вычислительной техники для решения профессиональных задач, в том числе с применением технологий искусственного интеллекта</p>	

4. Объем дисциплины (модуля) 1 з.е., в том числе 24 академических часа на контактную работу обучающихся с преподавателем, 12 академических часа на самостоятельную работу обучающихся.

5. Формат обучения очное обучение с использованием *дистанционных образовательных технологий*

6. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических или астрономических часов и виды учебных занятий

Наименование и краткое содержание разделов и тем дисциплины (модуля),  Форма промежуточной аттестации по дисциплине (модулю)	Всего (часы)	В том числе						
		Контактная работа (работа во взаимодействии с преподавателем) <i>Виды контактной работы, часы*</i>					Самостоятельная работа обучающегося <i>Виды самостоятельной работы, часы</i>	
		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Всего	Работа в среде электронного обучения	Всего
Тема 1. Основные понятия программирования	4	4				4	2	2
Тема 2. Язык программирования Python	4	4				4	2	2
Тема 3. Основные источники и форматы хранения данных	4	4				4	2	2
Тема 4. Алгоритмы обработки и анализа данных	6	6				6	3	3
Тема 5. Технологии искусственного интеллекта	6	6				6	3	3
<b>Итого:</b>	36	24				24	12	12

\* Текущий контроль успеваемости может быть реализован в рамках занятий семинарского типа, групповых или индивидуальных консультаций

\*\* Практическая подготовка (при наличии) осуществляется на базе ... (указать – структурное подразделение МГУ или организацию (предприятие), практическая подготовка на базе которого осуществляется на основании Договора)

\*\*\* Часы на проведение промежуточной аттестации выделяются из часов самостоятельной работы обучающегося.

7. Фонд оценочных средств для оценивания результатов обучения по дисциплине (модулю)

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	СРЕДСТВА ОЦЕНИВАНИЯ	ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю)			
		<i>Шкалы и критерии оценивания могут быть сформулированы как общие для всех дисциплин (модулей) и размещены в документе «Оценочные и методические материалы для контроля формирования компетенций у обучающихся в процессе освоения образовательной программы», входящем в состав ОПОП</i>			
		2 (не зачтено)	3 (зачтено)	4 (зачтено)	5 (зачтено)
<b>Знать:</b>					
Основные понятия программирования	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Этапы решения задач и применением средств вычислительной техники	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Виды и форматы хранения данных	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Этапы проведения анализа данных	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Стандартные алгоритмы обработки и	Тестирование	Отсутствие	Фрагментарные	Общие, но не	Сформированные

анализа данных		знаний	знания	структурированные знания	систематические знания
Основные понятия искусственного интеллекта и машинного обучения	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
<b>Уметь</b>					
Применять средства и технологии программирования для решения задач связанных с анализом данных	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Проводить поиск, обработку и анализ данных для решения профессиональных задач, в том числе открытых данных	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Использовать технологии искусственного интеллекта и машинного обучения для решения профессиональных задач	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
<b>Владеть</b>					
методами и технологиями использования средств вычислительной техники для решения профессиональных задач, в том числе с применением технологий искусственного интеллекта	Тестирование	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания

Типовые задания и иные материалы, необходимые для оценки результатов обучения

*Примерные тестовые задания*

Вопрос 1.

процесс выявления и исправления ошибок, несоответствий данных с целью улучшения их качества, иногда классифицируется как составная часть интеллектуального анализа данных.

+очистка данных

сбор данных  
анализ данных  
обогащение данных

### Вопрос 2.

процесс насыщения данных новой информацией, которая позволяет сделать их более ценными и значимыми с точки зрения решения той или иной задачи анализа.

очистка данных  
сбор данных  
анализ данных  
+обогащение данных

### Вопрос 3

Для каких из следующих задач может быть подходящим алгоритмом кластеризация методом K-средних?

1. Для множества электронных писем определить популярные темы. (+)
2. Получив набор новостных статей с различных новостных сайтов, выяснить являются ли они новостями культуры.
3. Для множества данных о растениях разделить их по типу соцветия. (+)
4. По профилю пользователя на сайте определить мужчина это или женщина.

### Вопрос 4

Степанида любит ходить на быстрые свидания. Однажды она задумалась о том насколько хорошо работает ее внутренний распознаватель качества потенциальных кавалеров. Она составила следующую таблицу (результаты за последний месяц):

	Предсказано «Плохой»	Предсказано «Хороший»
Плохой	96	4
Хороший	8	19

Рассчитайте какую долю хороших кавалеров из всех хороших кавалеров нашел алгоритм внутреннего распознавателя Степаниды.

0.704+  
0.900  
0.611  
0.809

### Вопрос 5

информация, размещаемая ее обладателями в сети "Интернет" в формате, допускающем автоматизированную обработку без предварительных изменений человеком в целях повторного ее использования

+Открытые данные

Набор открытых данных

Машиночитаемые данные

Метаданные

Вопрос 6

систематизированная совокупность открытых данных и метаинформации о них.

Открытые данные

+Набор открытых данных

Машиночитаемые данные

Метаданные

Вопрос 7

данные, представленные в форматах, пригодных для автоматического или автоматизированного использования.

Открытые данные

Набор открытых данных

+Машиночитаемые данные

Метаданные

Вопрос 8

Структурированные данные, представляющие собой характеристики описываемых сущностей для целей их идентификации, поиска, оценки, управления этими сущностями.

Открытые данные

Набор открытых данных

Машиночитаемые данные

+Метаданные

Вопрос 9

Среди приведенных языков укажите языки разметки

+XML

+HTML

PDF

CLIPS



## Вопрос 10

текстовый формат, предназначенный для представления табличных данных. Строка таблицы соответствует строке текста, которая содержит одно или несколько полей, разделенных запятыми

XML

HTML

+CSV

JSON

## Вопросы к зачету:

1. Перечислите области человеческой деятельности, где используются программные продукты.
2. Что такое программная система?
3. Что такое жизненный цикл программного обеспечения?
4. Перечислите этапы жизненного цикла программного обеспечения.
5. Перечислите основные подходы программирования.
6. Опишите структурный подход программирования.
7. Опишите процедурный подход программирования.
8. Опишите объектно-ориентированный подход программирования.
9. Дайте определение процедуры и функции.
10. Дайте определения программирования.
11. Дайте определение структуры данных.
12. Как Вы понимаете понятие «интерфейс».
13. Что такое поток управления и поток данных?
14. Проектирование интерфейса с пользователем. Структуры диалога; поддержка пользователя; многооконные интерфейсы; примеры реализации интерфейсов с пользователем с использованием графических пакетов.
15. Логические операции. Введение в pandas. Индексы в pandas.
16. Иерархические (множественные) индексы. Иерархические колонки.
17. Агрегирование в pandas. Объединение нескольких наборов данных
18. Машинное обучение. Дерево решений. Метод ближайших соседей.

19. Выбор параметров модели и кросс-валидация.
20. Практические примеры машинного обучения.

#### 8. Ресурсное обеспечение:

- Перечень основной и дополнительной учебной литературы
  - Основная литература
    1. Седжвик, Роберт, Программирование на Python: базовый курс : пер. с англ.. / Роберт Седжвик, Кевин Уэйн, Роберт Дондеро ; Принстон. ун-т. - СПб. : Диалектика, 2020. - 736 с.;
    2. Информатика. Базовый курс : [для бакалавров и специалистов] : учеб. пособие для студентов вузов. / под ред. С. В. Симоновича. - СПб. [и др.] : Питер, 2019. - 637, с
    3. Джоши, Пратик, Искусственный интеллект с примерами на Python : создание прил. искусств. интеллекта с помощью Python для взаимодействия с окружающим миром : [пер. с англ.]. / Пратик Джоши. - СПб. : Диалектика, 2019. - 444 с.;
  - Дополнительная литература
    1. Милехина, О. В. Информационные системы: теоретические предпосылки к построению : учебное пособие / О.В. Милехина, Е.Я. Захарова, В.А. Титова. - Новосибирск : НГТУ, 2014. - 383 с. - ISBN 978-5-7782-2405-6 : Б. ц.  
URL: [http://biblioclub.ru/index.php?page=book\\_view&book\\_id=258420](http://biblioclub.ru/index.php?page=book_view&book_id=258420)
    2. Жданов, С. А. Информационные системы : учебник / С.А. Жданов, М.Л. Соболева, А.С. Алфимова. - М. : Прометей, 2015. - 302 с. - ISBN 978-5-9906-2644-7 : Б. ц.  
URL: [http://biblioclub.ru/index.php?page=book\\_view&book\\_id=426722](http://biblioclub.ru/index.php?page=book_view&book_id=426722)
    3. Абрамов, Г. В. Проектирование информационных систем : учебное пособие / Г.В. Абрамов, И.Е. Медведкова, Л.А. Коробова. - Воронеж : ВГУИТ, 2012. - 172 с. - ISBN 978-5-89448-953-7 : Б. ц.  
URL: [http://biblioclub.ru/index.php?page=book\\_view\\_red&book\\_id=141626](http://biblioclub.ru/index.php?page=book_view_red&book_id=141626)
- Перечень лицензионного программного обеспечения
  - Язык программирования Python
- Перечень профессиональных баз данных и информационных справочных систем
  1. Система федеральных образовательных порталов. Информационно-коммуникационные технологии в образовании. <http://www.ict.edu.ru/lib/>
  2. Интернет университет информационных технологий. <http://www.intuit.ru/>

3. Система федеральных образовательных порталов. Информационно-коммуникационные технологии в образовании.  
<http://www.ict.edu.ru/lib/>

4. Российская национальная библиотека (РНБ). [www. hbl-russia.ru](http://www.hbl-russia.ru) <http://www. nlr. ru>.

5. Российская государственная библиотека (РГБ). <http://www. rsl. ru>.

6. ЭБС « Университетская библиотека онлайн» <http://www.biblioclub.ru>

7. ЭБС «Znanium.com» <http://znanium.com/>

8. ЭБС «Юрайт» <https://biblio-online.ru/>

- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)
- Описание материально-технической базы.

9. Язык преподавания.

Русский

10. Преподаватель (преподаватели).

11. Разработчики программы.

Якушин А.В., к.п.н., доцент, факультет ВМК МГУ