

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт биологии и биомедицины

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Основы систем искусственного интеллекта и анализа данных

Уровень высшего образования

Бакалавриат

Направление подготовки / специальность

06.03.01 - Биология

Направленность образовательной программы

Биология (общий профиль)

Форма обучения

очная

г. Нижний Новгород

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.ДВ.07.07 Основы систем искусственного интеллекта и анализа данных относится к части, формируемой участниками образовательных отношений образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства	
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	Для текущего контроля успеваемости	Для промежуточной аттестации
ПК-1: Способен осуществлять информационный поиск по выбранной научной тематике в области биологии, излагать и критически анализировать получаемую информацию, представлять результаты исследований в виде презентаций, научно-технических отчетов, обзоров, аналитических карт, пояснительных записок, публикаций в научных изданиях; поддерживать дискуссию по актуальным вопросам биологии и экологии	ПК-1.1: Знает: - правила сбора и анализа информации по теме исследования, способы и правила представления результатов в письменной и устной формах ПК-1.2: Умеет: - планировать и осуществлять поиск научной информации, оформлять результаты исследования для представления в письменной и устной формах ПК-1.3: Владеет: - опытом поиска, анализа, представления и обсуждения результатов исследования	ПК-1.1: Знает теоретические основы и современные достижения в области систем искусственного интеллекта и анализа данных ПК-1.2: Умеет применять экспериментальные и теоретические методы для исследований в области систем искусственного интеллекта и анализа данных ПК-1.3: Владеет устойчивыми навыками планирования и проведения исследования в системе искусственного интеллекта и анализа данных	Опрос Практическое задание Тест	Зачёт: Контрольные вопросы Практическое задание

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	2
Часов по учебному плану	72
в том числе	

аудиторные занятия (контактная работа):	
- занятия лекционного типа	24
- занятия семинарского типа (практические занятия / лабораторные работы)	24
- КСР	1
самостоятельная работа	23
Промежуточная аттестация	0 Зачёт

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

Наименование разделов и тем дисциплины	Всего (часы)	в том числе			
		Контактная работа (работа во взаимодействии с преподавателем), часы из них			Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа (практические занятия/ лабора- торные работы), часы	Всего	
	о ф о	о ф о	о ф о	о ф о	о ф о
История, основные тенденции, понятия и проблемы в области разработки систем искусственного интеллекта. Основные стандартные типы задач машинного обучения. Препроцессинг экспериментальных данных в библиотеке Pandas языка программирования Python	30	12	12	24	6
Логистическая и линейная регрессия.	14	4	4	8	6
Решающие деревья и нейронные сети	14	4	4	8	6
Кластерный анализ и метод главных компонентов	13	4	4	8	5
Аттестация	0				
КСР	1				1
Итого	72	24	24	49	23

Содержание разделов и тем дисциплины

Практические занятия /лабораторные работы организуются, в том числе, в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

На проведение практических занятий / лабораторных работ в форме практической подготовки отводится: очная форма обучения - 14 ч.

Практические занятия /лабораторные работы организуются, в том числе, в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

На проведение практических занятий / лабораторных работ в форме практической подготовки отводится: очная форма обучения - 14 ч.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведенным в п. 5.

Самостоятельная работа подразумевает работу с научной и научно-методической литературой.

Виды самостоятельной работы студентов в рамках освоения дисциплины:

- изучение понятийного аппарата и проработка тем дисциплины;
- работа с основной и дополнительной литературой дома и в библиотеке;
- изучение сайтов по темам дисциплины в сети Интернет;
- подготовка к занятиям практического типа (устный опрос и работа с дистрибутивом Anaconda языка программирования Python);
- подготовка к тестам;
- подготовка к зачету.

5. Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

5.1 Типовые задания, необходимые для оценки результатов обучения при проведении текущего контроля успеваемости с указанием критериев их оценивания:

5.1.1 Типовые задания (оценочное средство - Опрос) для оценки сформированности компетенции ПК-1:

1. Назовите основные этапы развития области искусственного интеллекта?
2. Какие проблемы присутствуют при разработке систем искусственного интеллекта?
3. Что такое машинное обучение? Назовите основные решаемые задачи?
4. Что такое предпроцессинг экспериментальных данных?
1. Что такое логистическая регрессия?
2. Что такое линейная регрессия?
3. В чем отличие логистической регрессии от линейной?
4. Что такое решающие деревья?
5. Что такое перцептрон?
6. Структура и модель искусственного нейрона?
7. Какие функции активации бывают?
8. Что такое кластерный анализ данных?
9. Что такое метод главных компонент?
10. Для чего применяется метод главных компонент?

Критерии оценивания (оценочное средство - Опрос)

Оценка	Критерии оценивания
зачтено	50-100% правильных ответов
не зачтено	менее 50% правильных ответов

5.1.2 Типовые задания (оценочное средство - Практическое задание) для оценки сформированности компетенции ПК-1:

1. Создайте dataframe с именем my_data, в котором две колонки с именами (type - строки,

value - целые числа) и четыре наблюдения в каждой колонке: type value

A 10

1. 14

2. 12

B 23.

Необходимо написать код программы.

1. В dataframe с именем my_stat сохранено 20 строк и четыре колонки (V1, V2, V3, V4). Напишите код, который позволит сохранить в следующие переменные части датафрейма:

- В переменную с именем subset_1 сохраните только первые 10 строк и только 1 и 3 колонку.
- В переменную с именем subset_2 сохраните все строки кроме 1 и 5 и только 2 и 4 колонку.

Критерии оценивания (оценочное средство - Практическое задание)

Оценка	Критерии оценивания
зачтено	Задание выполнено, допускаются незначительные недочеты
не зачтено	Задание не выполнено

5.1.3 Типовые задания (оценочное средство - Тест) для оценки сформированности компетенции ПК-1:

1. Какие способы позволяют отобрать 5 первых строк датафрейма df:

1.df.loc[:5] 2. df.tail(5) 3. df.iloc[0:5] 4. df.loc[:5]

2. Какой метод возвращает серию с описанием типа каждой колонки: 1.df.size 2. df.columns 3. df.dtypes 4. df.index

3. Выберите правильное утверждение по методу filter:

1. Можно использовать только один из аргументов items, like, regex 2. Можно использовать только два из аргументов items, like, regex 3. Можно использовать одновременно все аргументы items, like, regex

4. Аргумент items:

1. отбирает лэйблы, переданные в коллекции в качестве значения 2. отбирает те лэйблы, где встречается поданное значение 3. отбирает лэйблы, подходящие к переданному паттерну 4. отвечает за то, какие лэйблы мы рассматриваем - индекс или колонки

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
зачтено	50-100% правильных ответов
не зачтено	менее 50% правильных ответов

5.2. Описание шкал оценивания результатов обучения по дисциплине при промежуточной аттестации

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатора достижения компетенций)	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено			зачтено			
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено много негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками.	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены	Продemonстрированы все основные умения. Решены все основные задачи с	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в

	ответа		Выполнены все задания, но не в полном объеме	ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	все задания в полном объеме, но некоторые с недочетами	отдельным и несущественными недочетами, выполнены все задания в полном объеме	полном объеме без недочетов
<u>Навыки</u>	Отсутствие базовых навыков. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов	Продемонстрирован творческий подход к решению нестандартных задач

Шкала оценивания при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно», продемонстрированы знания, умения, владения по соответствующим компетенциям на уровне выше предусмотренного программой
	отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично».
	очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо»
	хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо».
	удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно».
	плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.3 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения на промежуточной аттестации с указанием критериев их оценивания:

5.3.1 Типовые задания (оценочное средство - Контрольные вопросы) для оценки сформированности компетенции ПК-1

1. Назовите основные проблемы в области разработки системы искусственного интеллекта?

Что из себя представляет процесс

разработки системы искусственного интеллекта?

2. Для чего необходим предпроцессинг данных? Что входит предпроцессинг данных?

1. Что такое машинное обучение? Какие типы задач есть в

2. 2. машинном

обучении и какими методами они решаются?

4. Дайте определение регрессионному анализу? Дайте определение принцип линейной регрессии? Приведите пример задачи.

5. Дайте определение и принцип логистической регрессии? В чем отличие логистической регрессии от линейной регрессии? Приведите пример задачи.

6. Дайте определение и принцип метода решающих деревьев.

Приведите пример задачи и основные ограничения метода.

Дайте определение и принцип метода нейронных сетей? Как методы обучения нейронных сетей бывают? Что такое перцептрон .

Приведите пример задачи

Что такое и принцип метода главных компонент? Для чего используется? Приведите пример задачи

Что такое и принцип метода кластерного анализа? Для чего используется? Приведите пример задачи

Критерии оценивания (оценочное средство - Контрольные вопросы)

Оценка	Критерии оценивания
зачтено	50-100% правильных ответов
не зачтено	менее 50% правильных ответов

5.3.2 Типовые задания (оценочное средство - Практическое задание) для оценки сформированности компетенции ПК-1

1. Создайте dataframe с именем `my_data`, в котором две колонки с именами (`type` - строки, `value` - целые числа) и четыре наблюдения в каждой колонке: `type value`

A 10

A 14

B 12

B 23.

Необходимо написать код программы.

2. В dataframe с именем `my_stat` сохранено 20 строк и четыре колонки (`V1`, `V2`, `V3`, `V4`). Напишите код, который позволит сохранить в следующие переменные части датафрейма:

✓ В переменную с именем `subset_1` сохраните только первые 10 строк и только 1 и 3 колонку.

✓ В переменную с именем `subset_2` сохраните все строки кроме 1 и 5 и только 2 и 4 колонку.

3. Необходимо отобрать данные из dataframe с именем `my_stat` четыре колонки `V1`, `V2`, `V3`, `V4` и записать в следующие переменные:

✓ В переменную `subset_1` сохраните только те наблюдения, у которых значения переменной `V1` строго больше 0, и значение переменной `V3` равняется 'A'.

✓ В переменную `subset_2` сохраните только те наблюдения, у которых значения переменной `V2` не равняются 10, или значения переменной `V4` больше или равно 1. Необходимо написать код программы.

4. В датафрейме `my_stat` с четырьмя колонками (`V1`, `V2`, `V3`, `V4`) и 20 строками необходимо создать две переменные:

✓ $V5 = V1 + V4$

✓ $V6 = \ln(V2)$ (натуральный логарифм переменной `V2`). Необходимо написать код программы.

5. Подсчитайте энтропию по признакам в группах, где признак равен 0 и 1 соответственно для следующих данных:

№	Гавкает	Лазает по деревьям	Вид
0	1	0	Собачка
1	1	0	Собачка
2	1	0	Собачка
3	1	0	Собачка
4	0	1	Котик
5	0	1	Котик
6	0	1	Котик
7	0	1	Котик
8	1	1	Котик
9	0	1	Котик

6. Подсчитайте IG (information Gain) для признака гавкает по следующим данным:

№	Гавкает	Лазает по деревьям	Вид
0	1	0	Собачка
1	1	0	Собачка
2	1	0	Собачка
3	1	0	Собачка
4	0	1	Котик
5	0	1	Котик
6	0	1	Котик
7	0	1	Котик
8	1	1	Котик
9	0	1	Котик

7. Подсчитайте IG (information Gain) для признака лазает по деревьям по следующим данным:

№	Гавкает	Лазает по деревьям	Вид
0	1	0	Собачка
1	1	0	Собачка
2	1	0	Собачка
3	1	0	Собачка
4	0	1	Котик
5	0	1	Котик
6	0	1	Котик
7	0	1	Котик
8	1	1	Котик
9	0	1	Котик

8. Подсчитайте IG (information Gain) для признака лазает по деревьям по следующим данным:

№	Гавкает	Лазает по деревьям	Вид
0	1	0	Собачка
1	1	0	Собачка
2	1	0	Собачка
3	1	0	Собачка
4	0	1	Котик
5	0	1	Котик
6	0	1	Котик
7	0	1	Котик

1. Гудфеллоу Я. Глубокое обучение : монография / Гудфеллоу Я.; Бенджио И.; Курвилль А. - Москва : ДМК-пресс, 2018. - 652 с. - ISBN 978-5-97060-618-6.,
<https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=772962&idb=0>.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

<http://www.ncbi.nlm.nih.gov/pubmed> [webofknowledge.com](http://www.ncbi.nlm.nih.gov/pubmed)

www.scopus.com [elsevierscience.ru](http://www.elsevierscience.ru) elibrary.ru Wikipedia.ru

Профессиональный ресурс по машинному обучению <https://machinelearning.ru>

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению подготовки/специальности 06.03.01 - Биология.

Автор(ы): Стасенко Сергей Викторович, кандидат физико-математических наук.

Заведующий кафедрой: Казанцев Виктор Борисович, доктор физико-математических наук.

Программа одобрена на заседании методической комиссии от 05.12.2023 г., протокол № 2.