

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Дзержинский филиал ННГУ

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Введение в анализ данных и искусственный интеллект

Уровень высшего образования

Бакалавриат

Направление подготовки / специальность

44.03.05 - Педагогическое образование (с двумя профилями подготовки)

Направленность образовательной программы

Начальное образование. Психолого-педагогическое сопровождение
образовательного процесса

Форма обучения

очная, заочная

г. Дзержинск

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина ФТД.01 Введение в анализ данных и искусственный интеллект является факультативом в образовательной программе.

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
ОПК-9: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-9.1: Знает принципы работы современных информационных технологий ОПК-9.2: Умеет осуществлять отбор современных информационных технологий для решения стандартных задач профессиональной деятельности ОПК-9.3: Владеет навыками использования современных информационных технологий для решения задач профессиональной деятельности	ОПК-9.1: Знать: принцип, порядок поиска и анализа и отбора информации по использованию программного средства, выбора и использования необходимых функций программных средств для решения конкретной задачи. ОПК-9.2: Уметь: находить и анализировать новые знания по использованию программного средства, выбирать и использовать необходимые функции программных средств для решения конкретной задачи ОПК-9.3: Владеть: навыками применения современных информационных технологий и классификации программных средств для решения практических задач, в том числе отечественного производства при решении задач профессиональной деятельности.	Практическое задание Собеседование Тест	Зачёт: Контрольные вопросы

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная	заочная
Общая трудоемкость, з.е.	1	1
Часов по учебному плану	36	36
в том числе		
аудиторные занятия (контактная работа):		
- занятия лекционного типа	8	8
- занятия семинарского типа (практические занятия / лабораторные работы)	8	8
- КСР	1	1
самостоятельная работа	19	15
Промежуточная аттестация	0 Зачёт	4 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/лабораторные работы), часы		Всего				
	0 Ф 0	3 Ф 0	0 Ф 0	3 Ф 0	0 Ф 0	3 Ф 0	0 Ф 0	3 Ф 0	0 Ф 0	3 Ф 0	
Тема 1. Постановки и примеры задач	5	5	2	2			2	2	3	3	
Тема 2. Введение в язык Python	6	5			2	2	2	2	4	3	
Тема 3. Описательная статистика и разведочный анализ данных	8	7	2	2	2	2	4	4	4	3	
Тема 4. Задачи классификации и регрессии	8	7	2	2	2	2	4	4	4	3	
Тема 5. Задачи обучения без учителя	8	7	2	2	2	2	4	4	4	3	
Аттестация	0	4									
КСР	1	1						1	1		
Итого	36	36	8	8	8	8	17	17	19	15	

Содержание разделов и тем дисциплины

Тема 1. Постановки и примеры задач

Тема 2. Введение в язык Python

Тема 3. Описательная статистика и разведочный анализ данных

Тема 4. Задачи классификации и регрессии

Тема № 5 Задачи обучения без учителя

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 студента при изучении дисциплины «Введение в анализ данных и искусственный интеллект» включает выполнение практических заданий под контролем преподавателя и подготовку к зачету.

Контрольные и тестовые вопросы, практические задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

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 - Практическое задание) для оценки сформированности компетенции ОПК-9:

1. Эссе “Мое любимое приложение, использующее технологию искусственного интеллекта”
2. Практическое задание “Программа на Python “Угадай число”.
 - а. Написать программу, которая умеет играть с человеком в игру “Угадай число”. Программа загадывает случайно число от 1 до 100. Человек должен угадать его. Для этого он делает несколько (не более 7) попыток, называя какое-то число в этом диапазоне. На каждую попытку человека компьютер говорит, угадал человек или нет, и в случае, если человек не угадал, больше или меньше загаданное число числа, названного человеком. Напишите такую программу. Для ввода информации человеком используйте функцию input(). Для вывода - print().
 - б. Реализуйте программу, играющую в игру "Угадай число", но человек и компьютер меняются ролями.
3. Практическое задание “Разведочный анализ данных”.
 1. Найдите интересные данные, с которыми вы хотите работать. Рекомендуемые ресурсы: <https://www.kaggle.com/datasets>, <https://www.kaggle.com/competitions>, <https://archive.ics.uci.edu/ml/index.php>. Рекомендуется работать с табличными данными в несколько десятков признаков (столбцов), несколько тысяч (строк).
 2. Опишите, что представляют из себя эти данные, какие признаки, почему они вам интересны.
 3. Примените к данным подходящие методы очистки (борьбы с выбросами и пропущенными значениями) и визуализации. Сделайте выводы.
4. Сделайте выводы

4. Практическое задание “Решение задачи классификации и регрессии”

- a. Для данных из задания 3 определите, какой признак вы будете определять по остальным. Что это за задача: классификации или регрессии?
- b. Подготовьте данные (отмасштабируйте количественные и бинаризируйте категориальные признаки).
- c. Разбейте выборку на обучающую и тестовую.
- d. Обучите несколько методов машинного обучения (не менее трех, например, kNN, RandomForests, MLP). Сравните качество этих методов на обучающей и тестовой выборках.
- e. Сделайте выводы

5. Практическое задание “Решение задачи кластеризации”.

- a. Для данных из задания 3 сформулируйте задачу кластеризации.
- b. Обучите несколько методов кластеризации (не менее двух, например, k-means, DBSCAN). Объясните ваш выбор. Сравните результаты работы алгоритмов, а также полученные результаты с результатами работы методов обучения с учителем из задания 4.
- c. Сделайте выводы

Критерии оценивания (оценочное средство - Практическое задание)

Оценка	Критерии оценивания
превосходно	Задание выполнено в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, обучающийся отвечает четко и последовательно, показывает глубокое знание основного и дополнительного материала
отлично	Задание выполнено в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, обучающийся отвечает четко и последовательно, показывает глубокое знание основного материала
очень хорошо	Задание выполнено в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, обучающийся отвечает четко и последовательно, показывает глубокое знание материала, допущено не более 2 неточностей не принципиального характера
хорошо	Задание выполнено в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, допущены неточности не принципиального характера, но обучающийся показывает систему знаний по теме своими

Оценка	Критерии оценивания
	ответами на поставленные вопросы
удовлетворительно	Задание выполнено не в полном объеме (решено более 50% поставленных задач), но обучающийся допускает ошибки, нарушена последовательность ответа, но в целом раскрывает содержание основного материала
неудовлетворительно	Задание выполнено не в полном объеме (решено менее 50% поставленных задач), обучающийся дает неверную информацию при ответе на поставленные задачи, допускает грубые ошибки при толковании материала, демонстрирует незнание основных терминов и понятий.
плохо	Задание не выполнено, обучающийся демонстрирует полное незнание материала

5.1.2 Типовые задания (оценочное средство - Собеседование) для оценки сформированности компетенции ОПК-9:

На собеседовании проверяется, как студент освоил основные понятия. Он должен уметь давать развернутый ответ на следующие вопросы.

1. Что такое искусственный интеллект?
2. Что такое слабый (специализированный или прикладной) искусственный интеллект?
3. Что такое сильный искусственный интеллект?
4. Что такое глубокое обучение?
5. Что такое экспертная система?
6. Что такое машинное обучение?
7. Что такое анализ данных?
8. Что такое обучение с учителем?
9. Что такое обучающая выборка?
10. Что такое обучение без учителя?
11. Что такое классификация?
12. Что такое регрессия?
13. Что такое кластеризация?

14. Что такое иерархическая кластеризация?

15. Основные библиотеки и программные средства для решения задач анализа данных и машинного обучения.

Критерии оценивания (оценочное средство - Собеседование)

Оценка	Критерии оценивания
превосходно	Ответ дан в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, обучаемый отвечает четко и последовательно, показывает глубокое знание основного и дополнительного материала.
отлично	Ответ соответствует изучаемой теме, студент корректно отражает основные принципы и законы, учитывает в докладе особенности современного этапа развития науки, ссылается на авторитетных авторов, использует достоверные источники информации. Студент способен выразить идею ясно, удерживая внимание и интерес слушателя.
очень хорошо	Ответ дан в полном объеме (все поставленные задачи решены), ответ логичен и обоснован, обучаемый отвечает четко и последовательно, показывает глубокое знание материала, допущено не более 2 неточностей не принципиального характера
хорошо	Ответ соответствует изучаемой теме, студент корректно отражает основные принципы и законы, но при ответе не учитывает идеи современных авторов. Студент способен выразить свои идеи, но не может при этом удержать внимание и интерес слушателя.
удовлетворительно	Ответ соответствует изучаемой теме, но не в полной мере раскрывает ее, студент не ссылается на авторитетных авторов, использует недостоверные источники информации. раскрывает его содержание, не является удобной для восприятия. Студент с трудом выражает свою мысль, сталкивается со сложностями при ответе на дополнительные вопросы.
неудовлетворительно	Ответ не соответствует объявленной теме или не раскрывает ее содержания.
плохо	Ответ не дан, обучаемый демонстрирует полное незнание материала

5.1.3 Типовые задания (оценочное средство - Тест) для оценки сформированности компетенции ОПК-9:

. Что такое искусственный интеллект?

- Искусственный интеллект — это программа или устройство, имитирующее интеллектуальные способности человека.

- Искусственный интеллект — это кибер-механическое устройство, наполовину — человек, наполовину — машина. Под ним — боевое шасси из гиперсплава, управляемое микропроцессором, полностью бронированное, очень прочное. Но снаружи — живая человеческая ткань: плоть, кожа, волосы, кровь, выращенные для киборгов.

- Искусственный интеллект — это искусственная нейронная сеть, способная к обучению и самообучению.

2. Что такое слабый искусственный интеллект?

- Слабый искусственный интеллект - это программа, устройство или программно-аппаратный комплекс, умеющая решать только некоторый класс интеллектуальных задач.

- Слабый искусственный интеллект - это киборг модели от T-800 до T-888.

- Слабый искусственный интеллект - это однослойная искусственная нейронная сеть.

3. Что такое сильный искусственный интеллект?

- Сильный искусственный интеллект - это программа, устройство или программно-аппаратный комплекс, по своим интеллектуальным способностям ни в чем не уступающая человеку.

- Сильный искусственный интеллект - это киборг модели от T-900 и выше.

- Сильный искусственный интеллект - это любая глубокая нейронная сеть, умеющая решать задачи классификации и регрессии.

4. Что такое экспертная система?

- Экспертная система - это программная система позволяющая решать задачи искусственного интеллекта с помощью правил, сформулированных экспертом, составляющих базу знаний о предметной области.

- Экспертная система - это коллектив экспертов в данной предметной области, решающих интеллектуальные задачи.

- Экспертная система - это искусственная многослойная нейронная сеть.

5. Что такое машинное обучение?

- Машинное обучение - это подход в искусственном интеллекте, основанный на построении моделей (решающих правил) по данным.

- Машинное обучение - это система обучения при помощи информационных и электронных технологий.

- Машинное обучение - это целенаправленное объективное отражение действительности

6. Что такое анализ данных?

- Анализ данных - это извлечение закономерностей и знаний из данных.

- Анализ данных - это подход в машинном обучении, основанный на использовании глубоких нейронных сетей.

- Анализ данных - это извлечение закономерностей с помощью знаний, сформулированных экспертом.

7. Что такое обучающая выборка?

- Обучающая выборка - это набор объектов (входов) с известным для каждого объекта значением ответа (выхода, метки, целевой переменной).

- Обучающая выборка - это набор объектов, которые пользователь выбирает для своего обучения.

- Обучающая выборка - это набор объектов, которые человек удаляет из базы данных для эффективного обучения.

8. Что такое обучение с учителем?

- Обучение с учителем - это построение модели по данным.

- Обучение с учителем - это решение задач с помощью интеллектуального помощника.

- Обучение с учителем - это задача разбиения данных на группы похожих друг на друга объектов.

9. Что такое обучение без учителя?

- Обучение без учителя - это установление структурных свойств набора объектов.

- Обучение без учителя - это обучение человека или робота полезным навыкам без интеллектуальных помощников.

- Обучение без учителя - это задача определения по данным категориальных значений.

10. Что такое классификация?

- Классификация - это задача обучения с учителем, в которой ответ - это номер класса.
- Классификация - это задача определения по данным числовых значений.
- Классификация - это задача обучения без учителя, в которой необходимо разбить данные на группы похожих объектов.

11. Что такое регрессия?

- Регрессия - это задача обучения с учителем, в которой ответы принимают количественные (числовые) значения.
- Регрессия - это задача, в которой необходимо разбить множество объектов на группы похожих друг на друга объектов.
- Регрессия - это задача обучения с учителем, в которой необходимо определить по данным номер класса.

12. Что такое кластеризация?

- Кластеризация - это задача обучения без учителя, в которой необходимо разбить данные на группы похожих объектов.
- Кластеризация - это задача обучения с учителем, в которой необходимо определить по данным номер класса.
- Кластеризация - это задача обучения с учителем, в которой ответы принимают количественные значения.

13. Что такое иерархическая кластеризация?

- Иерархическая кластеризация - это задача обучения без учителя, в которой необходимо построить иерархию объектов, в которой вершине соответствует «кластер», содержащий все объекты, и на каждом уровне каждый кластер разбивается на два или более подкластеров.
- Иерархическая кластеризация - это задача классификации, в которой каждый объект может иметь несколько меток.
- Иерархическая кластеризация - это метод решения задачи кластеризации с помощью иерархии нейронных сетей.

Следующие вопросы предполагают ввод числа.

Ответ следует записать в скобках.

14. Рост детей в группе детского сада задается следующими значениями (в см.):

118, 112, 116, 114, 116, 115, 119, 111, 114, 117, 118

- Найдите медиану ()
- Найдите нижний квартиль ()
- Найдите верхний квартиль ()

15. Вес детей в группе детского сада задается значениями (в кг):

23, 23, 18, 18, 20, 19, 20, 17, 22, 20, 20

- Вычислите среднее ()
- Вычислите несмещенную оценку дисперсии s^2 ()
- На основе несмещенной оценки дисперсии вычислите оценку среднеквадратического отклонения s ()

16. На тестовой выборке, содержащей 2000 изображений котиков и собачек (950 котиков, 1050 собачек), нейронная сеть совершила 202 ошибки: 76 котиков были классифицированы как собачки, а 126 собачек – как котики. Считая положительным – класс собачек, вычислите метрики:

- Accuracy ()
- Error ()
- Precision ()
- Recall ()
- F1-мера ()

Ответ округлите до 3 правильных цифр после десятичной точки.

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
превосходно	100 % правильных ответов
отлично	90-100% правильных ответов
очень хорошо	70-89% правильных ответов
хорошо	60-69 % правильных ответов
удовлетворительно	50-60% правильных ответов
неудовлетворительно	40-50 % правильных ответов
плохо	Менее 40 % 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 ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с	Продемонстрированы все основные умения. Решены все основные задачи с отдельными несущественными недочетами	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов

			объеме	некоторые с недочетами	недочетами	и, выполнены все задания в полном объеме	
<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 ОПК-9

1. Что такое искусственный интеллект?
2. Что такое слабый (специализированный или прикладной) искусственный интеллект?
3. Что такое сильный искусственный интеллект?

4. Что такое экспертная система?
5. Что такое машинное обучение?
6. Что такое анализ данных?
7. Что такое обучение с учителем?
8. Что такое обучающая выборка?
9. Что такое обучение без учителя?
10. Что такое классификация?
11. Что такое регрессия?
12. Что такое кластеризация?
13. Что такое иерархическая кластеризация?

Критерии оценивания (оценочное средство - Контрольные вопросы)

Оценка	Критерии оценивания
зачтено	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно»
не зачтено	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно»

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Воронина В. В. Теория и практика машинного обучения : учебное пособие / Воронина В. В. - Ульяновск : УлГТУ, 2017. - 290 с. - Библиогр.: доступна в карточке книги, на сайте ЭБС Лань. - Книга из коллекции УлГТУ - Информатика. - ISBN 978-5-9795-1712-4., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=750984&idb=0>.

Дополнительная литература:

1. Басалин Павел Дмитриевич. Модели и методы интеллектуальной поддержки процессов принятия решений : учебно-методическое пособие / П. Д. Басалин, К. В. Безрук, М. В. Радаева ; ННГУ им. Н. И. Лобачевского. - Нижний Новгород : Изд-во ННГУ, 2018. - 134 с. - Текст : электронный., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=822376&idb=0>.

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

1. Python: <https://www.python.org/>
2. Anaconda: The Most Popular Python Data Science Platform
<https://www.anaconda.com/download/>
3. Google colab: colab.research.google.com

7. Материально-техническое обеспечение дисциплины (модуля)

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

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 44.03.05 - Педагогическое образование (с двумя профилями подготовки).

Автор(ы): Дюдяков Алексей Александрович, кандидат медицинских наук.

Программа одобрена на заседании методической комиссии от 23.12.2023, протокол № 17.