

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Радиофизический факультет

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Основы машинного обучения

Уровень высшего образования

Магистратура

Направление подготовки / специальность

02.04.02 - Фундаментальная информатика и информационные технологии

Направленность образовательной программы

Биоинформатика

Форма обучения

очная

г. Нижний Новгород

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.О.09 Основы машинного обучения относится к обязательной части 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
ОПК-2: Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	ОПК-2.1: Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ ОПК-2.2: Умеет анализировать типовые языки программирования, составлять программы ОПК-2.3: Имеет практический опыт решения задач анализа) интеграции различных типов программного обеспечения, анализа типов коммуникации	ОПК-2.1: Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ ОПК-2.2: Умеет анализировать типовые языки программирования, составлять программы ОПК-2.3: Имеет практический опыт решения задач анализа) интеграции различных типов программного обеспечения, анализа типов коммуникации	Контрольная работа	Зачёт: Проект
ОПК-3: Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и	ОПК-3.1: Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей ОПК-3.2: Умеет	ОПК-3.1: Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей	Контрольная работа	Зачёт: Проект

математического моделирования	соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, образовательного контента, средств тестирования систем ОПК-3.3: Имеет практический опыт применения разработки программного обеспечения и тестирования программных продуктов	ОПК-3.2: Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, образовательного контента, средств тестирования систем ОПК-3.3: Имеет практический опыт применения разработки программного обеспечения и тестирования программных продуктов		
-------------------------------	--	---	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	3
Часов по учебному плану	108
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	32
- занятия семинарского типа (практические занятия / лабораторные работы)	0
- КСР	1
самостоятельная работа	75
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

Введение в анализ данных	10	4		4	6
Обзор пакетов Python	8	2		2	6
линейная регрессия	14	4		4	10
Логистическая регрессия	11	2		2	9
Анализ временных рядов	16	6		6	10
Градиентный бустинг, деревья решений	14	4		4	10
Кластерный анализ	18	6		6	12
Введение в нейронные сети	16	4		4	12
Аттестация	0				
КСР	1			1	
Итого	108	32	0	33	75

Содержание разделов и тем дисциплины

Понятия генеральной и выборочной совокупности. Принципы статистического оценивания параметров. Модели сигналов и шумов. Точечные оценки неизвестных параметров. Априорные и апостериорные плотности вероятности оцениваемых параметров. Функции потерь, средние потери, критерии оптимальности оценок. Методы нахождения точечных оценок. Метод моментов. Метод максимального правдоподобия. Свойства точечных оценок параметров. Интервальные оценки неизвестных параметров. Доверительные интервалы для параметров нормального распределения. Доверительные интервалы в случае асимптотически нормальных оценок. Определение необходимого объема выборки. Примеры оценки параметров в пакетах прикладных программ.

2. Пакеты Pandas, Numpy, Statistics

3. Линейная регрессионная модель

Линейная регрессионная модель и ее приложения. Оценивание коэффициентов регрессии. Связь метода наименьших квадратов и метода максимального правдоподобия. Оценка дисперсии ошибок. Оценка ковариационной матрицы выборочных коэффициентов регрессии. Свойства коэффициентов при регрессорах. Дисперсионный анализ регрессионной модели. Коэффициент детерминации как характеристика качества модели. Проверка гипотезы о нормальном распределении остатков модели. Проверка гипотез о значимости коэффициентов регрессии. Доверительные интервалы для оценок параметров. Проверка адекватности регрессии. Примеры моделей регрессий в пакетах прикладных программ.

4. Логистическая регрессия, область применения, балансировка выборки, критерии оценки качества моделей

5. анализ одномерных временных рядов

Простейшие примеры стационарных и нестационарных случайных временных рядов (белый шум, временной ряд с линейным трендом, случайное блуждание и др.). Основные составляющие временного ряда. Лаговый и разностный операторы: определение и основные свойства. Исследование и моделирование неслучайной составляющей временного ряда. Основные типы трендов (детерминированный и стохастический). Критерии проверки наличия тренда во временных рядах; методы выделения тренда. Тест Чоу для анализа временного ряда. Основные понятия о методах скользящего среднего как методах выделения неслучайной составляющей.

6. Градиентный бустинг, Регрессионные деревья решений

7. Сущность кластерного анализа. Основные соотношения кластерного анализа. Процедуры кластерного анализа. Пример выполнения основных процедур кластерного анализа. Основные понятия кластерного

анализа. Расстояния между объектами. Меры близости между кластерами. Основные типы задач и алгоритмов кластерного анализа. Метод k средних. Иерархическая кластеризация. Использование статистических пакетов.

8. Структура нейронных сетей. Обзор пакетов Python. Критерии оценки качества моделей

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.

Платонов А. В. Машинное обучение : учебное пособие / А. В. Платонов. - Москва : Юрайт, 2023. - 85 с. - (Высшее образование). - ISBN 978-5-534-15561-7. - Текст : электронный // ЭБС "Юрайт".

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 - Контрольная работа) для оценки сформированности компетенции ОПК-2:

1. Основные пакеты Python, используемые для моделей машинного обучения.
2. Методы работы с пропущенными данными.
3. Фильтры для обработки изображений.

5.1.2 Типовые задания (оценочное средство - Контрольная работа) для оценки сформированности компетенции ОПК-3:

1. Применение деревьев решений.
2. Методы работы с текстовыми данными.

3. Пакеты работы с временными рядами.

Критерии оценивания (оценочное средство - Контрольная работа)

Оценка	Критерии оценивания
зачтено	Применяет методы анализа данных, строит модели, выполняет сравнительный анализ моделей
не зачтено	Не знает методы анализа данных, не строит модели, не выполняет сравнительный анализ моделей

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

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатор 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	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами	Продemonстрированы все основные умения. Решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в 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 некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов	Продemonстрирован 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 - Проект) для оценки сформированности компетенции ОПК-2

1. Построение кластеров на моделированных данных с применением деревьев решений.
2. Построение регрессии на предложенном датасете.
3. Построение логистической регрессии на выбранном датасете.
4. VAR модели временных рядов.

5.3.2 Типовые задания (оценочное средство - Проект) для оценки сформированности компетенции ОПК-3

1. Построение регрессии на предложенном датасете.
2. Построение логистической регрессии на выбранном датасете.
3. VAR модели временных рядов.
4. GARCH модели для оценки временных рядов.

Критерии оценивания (оценочное средство - Проект)

Оценка	Критерии оценивания
зачтено	Применяет методы анализа данных, строит модели, выполняет сравнительный анализ моделей
не зачтено	Не знает методы анализа данных, не строит модели, не выполняет сравнительный анализ моделей

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

Основная литература:

1. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения / Рашка С. - Москва : ДМК-пресс, 2017., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=659331&idb=0>.
2. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных : монография / Флах П. - Москва : ДМК-пресс, 2023. - 401 с. - ISBN 978-5-89818-300-4., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=878863&idb=0>.

Дополнительная литература:

1. Шарден Б. Крупномасштабное машинное обучение вместе с Python : монография / Шарден Б.; Массарон Л.; Боскетти А. - Москва : ДМК-пресс, 2018. - 358 с. - ISBN 978-5-97060-506-6., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=772971&idb=0>.
2. Платонов А. В. Машинное обучение : учебное пособие / А. В. Платонов. - Москва : Юрайт, 2023. - 85 с. - (Высшее образование). - ISBN 978-5-534-15561-7. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=841919&idb=0>.

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

Anaconda, Python 3,7 и выше

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 02.04.02 - Фундаментальная информатика и информационные технологии.

Автор(ы): Лапинова Светлана Александровна, кандидат физико-математических наук.

Заведующий кафедрой: Павлов Игорь Сергеевич, доктор физико-математических наук.

Программа одобрена на заседании методической комиссии от 18 декабря 2023, протокол № 09/23.