

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Теория машинного обучения

Уровень высшего образования

Магистратура

Направление подготовки / специальность

02.04.02 - Фундаментальная информатика и информационные технологии

Направленность образовательной программы

Искусственный интеллект

Форма обучения

очная

г. Нижний Новгород

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.ДВ.06.02 Теория машинного обучения 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-8: Способен к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности	ПК-8.1: Знать основы ИТ в области КС и иметь навыки анализа современного состояния науки и информационных технологий в области КС ПК-8.2: Иметь навыки проектирования и разработки и развития ИТ-решений на основе анализа современного состояния науки и информационных технологий в области КС ПК-8.3: Иметь навыки управления разработкой и развитием ИТ-решений на основе анализа современного состояния науки и информационных технологий в области КС	ПК-8.1: Знать теоретические основы машинного обучения, основные алгоритмы решения задачи восстановления регрессии, основные алгоритмы решения задачи классификации. ПК-8.2: Уметь профессионально разрабатывать и проводить процедуры тестирования алгоритмов кластеризации ПК-8.3: Владеть навыками решения практических задач с использованием методов машинного обучения; использования среды статистических вычислений R или библиотеки ScikitLearn для решения задач машинного обучения.	Задания	Зачёт: Контрольные вопросы

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	16
- занятия семинарского типа (практические занятия / лабораторные работы)	16
- 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/лабораторные работы), часы	Всего	
	о ф о	о ф о	о ф о	о ф о	о ф о
1. Введение.	19	2	2	4	15
2. Основы теории Вапника– Червоненкиса.	23	4	4	8	15
3. Основы теории PAC-обучения.	23	4	4	8	15
4. Лезвие Оккама.	23	4	4	8	15
5. Точное обучение.	19	2	2	4	15
Аттестация	0				
КСР	1			1	
Итого	108	16	16	33	75

Содержание разделов и тем дисциплины

1. Введение. Основные задачи теории обучения. 2. Основы теории Вапника– Червоненкиса. Лемма Бернштейна– Чернова. Равномерная сходимост эмпирического риска к ожидаемому риску в конечном случае. Размерность Вапника–Червоненкиса. Примеры: пороговые функции, нейронные сети, деревья решений. ϵ - сеть. Теорема Вапника–Червоненкиса о равномерной сходимости эмпирического риска к ожидаемому в бесконечном случае. 3. Основы теории PAC-обучения. Определение «вероятно почти корректного» (PAC–probably approximately correct learning) обучения. Обучение конъюнкциям. Труднорешаемость задачи обучения 3-ДНФ формулам. Обучение 3-КНФ формулам. 4. Лезвие Оккама. Использование принципа лезвия Оккама при обучении конъюнкциям. Обучение деревьям решений. 5. Точное обучение. Обучение с помощью вопросов. Вопросы принадлежности и эквивалентности. Размерность наущения. Верхние и нижние оценки сложности обучения на основе размерности наущения.

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.

Для подготовки к практическим занятиям, самостоятельной работы и промежуточной аттестации рекомендуются электронные ресурсы <http://www.uic.unn.ru/~zny/ml/>

1. Воронцов К.В. Машинное обучение. Курс лекций. <http://www.machinelearning.ru>.

<http://www.intuit.ru/studies/courses/13844/1241/info>

2. Введение в аналитику больших массивов данных

<http://www.intuit.ru/studies/courses/12385/1181/info>

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 - Задания) для оценки сформированности компетенции ПК-8:

1. Функцию $f: \mathbf{R}^d \rightarrow \{0, 1\}$ назовем *ящиком*, если существуют вещественные числа $a_1, a_2, \dots, a_d, b_1, b_2, \dots, b_d$, такие, что $f(x) = 1$ тогда и только тогда, когда $a_i \leq x \leq b_i$ ($i = 1, 2, \dots, d$).
Найти функцию роста и размерность Вапника-Червоненкиса для класса всех ящиков.
Проиллюстрировать на этом примере лемму Зауэра.
2. Пусть T_h — множество всех функций $f: \mathbf{R}^d \rightarrow \{0, 1\}$, вычисляемых бинарными деревьями решений, высоты не выше h . Найти функцию роста и размерность Вапника-Червоненкиса для класса T_h . Проиллюстрировать на этом примере лемму Зауэра.

Критерии оценивания (оценочное средство - Задания)

Оценка	Критерии оценивания
зачтено	Приведен обоснованный развернутый ответ, задания выполнены в полном объеме, некоторые с недочетами
не зачтено	При выполнении заданий имеют место грубые ошибки или задания не выполнены.

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

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатор достижения компет	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				

компетенций)							
<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 и 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 - Контрольные вопросы) для оценки сформированности компетенции ПК-8

1.	Лемма Бернштейна–Чернова.
2.	Равномерная сходимость эмпирического риска к ожидаемому риску в конечном случае.
3.	Размерность Вапника–Червоненкиса класса пороговых функций.
4.	Размерность Вапника–Червоненкиса класса нейронных сетей.
5.	Размерность Вапника–Червоненкиса для класса деревьев решений.
6.	ϵ -сеть. Теорема Вапника–Червоненкиса о равномерной сходимости эмпирического риска к ожидаемому в бесконечном случае.
7.	Определение «вероятно почти корректного» (PAC–probably approximately correct learning) обучения.
8.	PAC-обучение. Обучение конъюнкциям.
9.	PAC-обучение. Труднорешаемость задачи обучения 3-ДНФ формулам.
10.	PAC-обучение. Обучение 3-КНФ формулам.
11.	Лезвие Оккама. Использование принципа лезвия Оккама при обучении конъюнкциям.
12.	Использование принципа лезвия Оккама при обучении деревьям решений.
13.	Обучение с помощью вопросов. Вопросы принадлежности и эквивалентности. Размерность научения.
14.	Верхние оценки сложности обучения с помощью вопросов принадлежности на основе размерности научения.

15.	Нижние оценки сложности обучения с помощью вопросов на основе размерности научения.

Критерии оценивания (оценочное средство - Контрольные вопросы)

Оценка	Критерии оценивания
зачтено	Владение основным и дополнительным материалом достаточное или с незначительными ошибками и погрешностями
не зачтено	владение материалом, необходимым по данному предмету, недостаточно. Работу за время семестра можно оценить как неудовлетворительную

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

Основная литература:

1. Платонов А. В. Машинное обучение : учебное пособие / А. В. Платонов. - Москва : Юрайт, 2022. - 85 с. - (Высшее образование). - URL: <https://urait.ru/bcode/508804> (дата обращения: 14.08.2022). - ISBN 978-5-534-15561-7 : 319.00. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=821948&idb=0>.

Дополнительная литература:

1. 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>.
 2. 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>.

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

1. The R Project for Statistical Computing <https://www.r-project>
 2. Welcome to Python.org <https://www.python.org/>
 3. scikit-learn: machine learning in Python scikit-learn.org/

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 02.04.02 - Фундаментальная информатика и информационные технологии.

Автор(ы): Золотых Николай Юрьевич, доктор физико-математических наук, доцент.

Заведующий кафедрой: Золотых Николай Юрьевич, доктор физико-математических наук.

Программа одобрена на заседании методической комиссии от 13.12.2023, протокол № 3.