

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION

**Federal State Autonomous Educational Institution of Higher Education
«National Research Lobachevsky State University of Nizhny Novgorod»**

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО
решением Ученого совета ННГУ
протокол № 10 от 02.12.2024 г.

Working programme of the discipline

Machine learning theory

Higher education level

Master degree

Area of study / speciality

02.04.02 - Fundamental Informatics and Information Technology

Focus /specialization of the study programme

Artificial Intelligence and Data Analysis

Mode of study

full-time

Nizhny Novgorod

Year of commencement of studies 2025

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.ДВ.02.02 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	<p>ПК-8.1: Знает методику разработки новых алгоритмических, методических и технологических решений</p> <p>ПК-8.2: Умеет применять полученные знания для разработки новых алгоритмических, методических и технологических решений</p> <p>ПК-8.3: Имеет практический опыт составления технического задания на разработку информационной системы</p>	<p>ПК-8.1: Знать теоретические основы машинного обучения, основные алгоритмы решения задачи восстановления регрессии, основные алгоритмы решения задачи классификации.</p> <p>ПК-8.2: Уметь профессионально разрабатывать и проводить процедуры тестирования алгоритмов кластеризации</p> <p>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.</p>	Задания	Зачёт: Контрольные вопросы

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	5
Часов по учебному плану	180
в том числе	

аудиторные занятия (контактная работа):	
- занятия лекционного типа	16
- занятия семинарского типа (практические занятия / лабораторные работы)	16
- КСР	1
самостоятельная работа	147
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	
0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0	
1. Introduction.	32	2	2	4	28
2. Fundamentals of the Vapnik-Chervonenkis theory.	38	4	4	8	30
3. Basics of PAC learning theory.	38	4	4	8	30
4. Occam's Razor.	38	4	4	8	30
5. Precise training.	33	2	2	4	29
Аттестация	0				
КСР	1			1	
Итого	180	16	16	33	147

Contents of sections and topics of the discipline

1. Introduction. Main tasks of learning theory. 2. Fundamentals of Vapnik–Chervonenkis theory. Bernstein–Chernoff lemma. Uniform convergence of empirical risk to expected risk in the finite case. Vapnik–Chervonenkis dimension. Examples: threshold functions, neural networks, decision trees. ϵ -network. Vapnik–Chervonenkis theorem on uniform convergence of empirical risk to expected risk in the infinite case. 3. Fundamentals of PAC learning theory. Definition of “probably approximately correct” (PAC) learning. Conjunction learning. Intractability of the problem of learning 3-DNF formulas. Learning 3-CNF formulas. 4. Occam’s razor. Using Occam’s razor in learning conjunctions. Learning decision trees. 5. Exact learning. Learning with questions. Membership and equivalence questions. Learning Dimensionality. Upper and lower bounds on the complexity of learning based on the learning dimension.

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.

To prepare for practical classes, independent work and intermediate certification, electronic resources are recommended <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. Assessment tools for ongoing monitoring of learning progress and interim certification in the discipline (module)

5.1 Model assignments required for assessment of learning outcomes during the ongoing monitoring of learning progress with the criteria for their assessment:

5.1.1 Model assignments (assessment tool - Assignments) to assess the development of the competency 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 Вашика–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 Вашика–Червоненкиса для класса T_h . Проиллюстрировать на этом примере лемму Зауэра.

Assessment criteria (assessment tool — Assignments)

Grade	Assessment criteria
pass	A well-founded, detailed answer was given, the tasks were completed in full, some with shortcomings
fail	When completing tasks, gross errors occur or tasks are not completed.

5.2. Description of scales for assessing learning outcomes in the discipline during interim certification

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 объеме	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения. Решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме	Продемонстрированы 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	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми 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

Scale of assessment for interim certification

Grade		Assessment criteria
pass	outstanding	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "outstanding", the knowledge and skills for the relevant competencies have been demonstrated at a level higher than the one set out in the programme.
	excellent	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "excellent",
	very good	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "very good",
	good	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "good",
	satisfactory	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "satisfactory", with at least one competency

		developed at the "satisfactory" level.
fail	unsatisfactory	At least one competency has been developed at the "unsatisfactory" level.
	poor	At least one competency has been developed at the "poor" level.

5.3 Model control assignments or other materials required to assess learning outcomes during the interim certification with the criteria for their assessment:

5.3.1 Model assignments (assessment tool - Control questions) to assess the development of the competency ПК-8

1. Bernstein–Chernov lemma.
2. Uniform convergence of empirical risk to expected risk in the finite case.
3. Vapnik–Chervonenkis dimension of a class of threshold functions.
4. Vapnik–Chervonenkis dimension of a class of neural networks.
5. Vapnik–Chervonenkis dimension for a class of decision trees.
6. ϵ -network. Vapnik–Chervonenkis theorem on uniform convergence of empirical risk to expected risk in the infinite case.
7. Definition of “probably approximately correct” (PAC) learning.
8. PAC learning. Conjunction learning.
9. PAC learning. Intractability of the problem of learning 3-DNF formulas.
10. PAC learning. Learning 3-CNF formulas.
11. Occam's Razor. Using Occam's Razor to learn conjunctions.
12. Using Occam's Razor to learn decision trees.
13. Learning with questions. Membership and equivalence questions. Learning dimension.
14. Upper bounds on the complexity of learning with membership questions based on the learning dimension.
15. Lower bounds on the complexity of learning with questions based on the learning dimension.

Assessment criteria (assessment tool — Control questions)

Grade	Assessment criteria
pass	Knowledge of basic and additional material is sufficient or with minor errors and inaccuracies
fail	knowledge of the material necessary for this subject is not enough. Work during the semester can be

Grade	Assessment criteria
	assessed as unsatisfactory

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 - Fundamental Informatics and Information Technology.

Авторы: Золотых Николай Юрьевич, доктор физико-математических наук, доцент.

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

Программа одобрена на заседании методической комиссии от 02.12.2024, протокол № 5.