

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
**Федеральное государственное автономное образовательное учреждение высшего
образования**

«Нижегородский государственный университет им. Н.И. Лобачевского»

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО
решением Ученого совета ННГУ
протокол от
«30» ноября 2022 г. № 13

Рабочая программа дисциплины
Машинное обучение

Уровень высшего образования
магистратура

Направление подготовки (специальность)
090404 Программная инженерия

Профиль подготовки (специализация)
Технологии цифровой трансформации

Форма обучения
очная

Нижний Новгород
2023

1. Место дисциплины (модуля) в структуре ОПОП.

Дисциплина «Б1.О.08, Машинное обучение» относится к обязательной части Блока 1 «Дисциплины (модули)» направления подготовки 09.04.04 «Программная инженерия» профиля подготовки «Технологии цифровой трансформации». Дисциплина преподается в 1 семестре. Трудоемкость дисциплины составляет 6 зачетных единиц, 216 час., экзамен.

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
1	Блок 1. Дисциплины (модули) Обязательная часть	Дисциплина «Б1.О.08, Машинное обучение» относится к обязательной части ООП направления подготовки 09.04.04 «Программная инженерия».

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

Формируемые компетенции	Планируемые результаты обучения по дисциплине (модулю), характеризующие этапы формирования компетенций	Наименование оценочного средства
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1. Знать методы критического анализа проблемных ситуаций.	Собеседование (экзамен)
	УК-1.3. Владеть основами системного подхода к анализу проблемных ситуаций.	
	УК-1.2. Уметь вырабатывать стратегию действий при возникновении критических ситуаций.	
ОПК-2 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	УК-1.3. Владеть основами системного подхода к анализу проблемных ситуаций.	Задача (практическое задание)
	ОПК-2.1. Знать современные интеллектуальные технологии для решения профессиональных задач;	Собеседование (экзамен)
	ОПК-2.2. Уметь обосновывать выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач	Задача (практическое задание)
	ОПК-2.3. Иметь навыки разработки оригинальных программных средств, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	Задача (практическое задание)

3. Структура и содержание дисциплины

Объем дисциплины (модуля) составляет

6 зачетных единиц, всего **216** час., из которых

66 часов составляет **контактная** работа обучающегося с преподавателем:

32 часа занятия лекционного типа,

32 часа занятия семинарского типа (семинары, лабораторные работы и т.п.),

2 часа мероприятия промежуточной аттестации

150 час. составляет **самостоятельная** работа обучающегося (в т.ч. включая 36 час. подготовки к экзамену).

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

№ п/п	Наименование и краткое содержание разделов и тем дисциплины (модуля), форма промежуточной аттестации по дисциплине (модулю)	Часов						
		Всего	В том числе					Самостоятельная работа обучающегося
			Контактная работа (работа во взаимодействии с преподавателем), часы из них					
			Занятия лекционного типа	Занятия семинарского	Занятия лабораторного типа	Консультации	Всего	
1.	Содержательная постановка задачи машинного обучения. Дедуктивное и индуктивное обучения. Признаковое описание объектов. Обучения с учителем. Решающая функция. Обобщающая способность решающей функции (проблема качества обучения). Задачи классификации и задачи восстановления регрессии. Обучение без учителя. Примеры практических задач.	14	2	2			4	10
2.	Вероятностная постановка задачи обучения по прецедентам. Принцип минимизации эмпирического риска. Байесовская теория решений. Принцип максимума апостериорной вероятности. Регрессионная функция. Байесов классификатор. Метод ближайшего соседа в задачах классификации и восстановления	14	2	2			4	10

	регрессии. Теорема об оценке риска в методе ближайшего соседа.							
3.	Экспериментальные методы оценки качества обучения. Разделение данных на обучающую и тестовую выборки. Метод скользящего контроля.	14	2	2			4	10
4.	Метод наименьших квадратов для решения задачи восстановления регрессии. Его вывод на основе метода максимального правдоподобия. Система нормальных уравнений.	14	2	2			4	10
5.	Проблема переобучения при решении задачи восстановления регрессии. Методы борьбы с переобучением: сокращение числа параметров, регуляризация (ридж-регрессия), метод лассо.	14	2	2			4	10
6.	Наивный байесовский классификатор.	14	2	2			4	10
7.	Линейный дискриминантный анализ. Квадратичный дискриминантный анализ.	13	2	2			4	9
8.	Логистическая регрессия. Обучение логистической регрессии	13	2	2			4	9
9.	Многослойные нейронные сети. Алгоритм обучения персептрона как метод стохастического градиентного спуска. Регуляризация как метод борьбы с переобучением. Понятие о глубоких нейронных сетях.	13	2	2			4	9
10.	Машина опорных векторов. Ядра и спрямляющие пространства.	13	2	2			4	9
11.	Деревья решений. Метод CART (classification and regression trees) для решения задач классификации и восстановления регрессии. Методы обработки пропущенных значений.	13	2	2			4	9
12.	Ансамбли решающих правил (классификаторов). Простое и взвешенное голосование. Баггинг. Алгоритм случайных деревьев (random forest) и его модификации.	13	2	2			4	9
13.	Бустинг. Алгоритм AdaBoost. Оценка ошибки предсказания. Бустинг и аддитивные модели. Градиентный бустинг деревьев решений.	13	2	2			4	9
14.	Обучение без учителя. Кластеризация. Метод центров тяжести. Метод медоидов. ЕМ-алгоритм.	13	2	2			4	9

15.	Метод главных компонент	13	2	2			4	9
16.	Иерархическая кластеризация. Агломеративные и разделяющие методы.	13	2	2			4	9
	В т.ч. текущий контроль	2					2	
	Итого:	216	32	32			66	150
	Промежуточная аттестация: экзамен							

Практические занятия (семинарские 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.

Практическая подготовка предусматривает: изучение методических материалов, подготовку к вопросам для собеседования, выполнение практических заданий и задач, выполнение заданий в форме вопросов к экзамену.

На проведение практических занятий (семинарских занятий /лабораторных работ) в форме практической подготовки отводится 32 часа.

Практическая подготовка направлена на формирование и развитие:

- практических навыков в соответствии с профилем ОП: создание и сопровождение архитектуры программных средств, разработка и тестирование программного обеспечения;
- компетенций – УК-1, ОПК-2.

Текущий контроль успеваемости проходит в рамках занятий семинарского и практического типа. Итоговый контроль (промежуточная аттестация) по дисциплине осуществляется на экзамене.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся осуществляется в виде работы с рекомендованной обязательной и дополнительной литературой, подготовке к лекциям, подготовке к экзамену и решения задач для самостоятельной работы.

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

Для обеспечения самостоятельной работы обучающихся используется открытый электронный онлайн-курс (Машинное обучение, <https://online.edu.ru/public/course?faces-redirect=true&cid=11092867>), созданный в системе открытых онлайн-курсов российских вузов <https://online.edu.ru/>.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине, включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности 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, без ошибок.	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 несущественным 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 владения материалом. 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 оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «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», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1 Контрольные вопросы к экзамену

Вопрос	Код формируемой компетенции
1) Вероятностная постановка задачи обучения по прецедентам. Принцип минимизации эмпирического риска. Байесовская теория решений. Принцип максимума апостериорной вероятности. Байесов классификатор.	УК-1
2) Экспериментальные методы оценки качества обучения.	УК-1
3) Метод наименьших квадратов для решения задачи восстановления регрессии. Его вывод на основе метода максимального правдоподобия. Линейная регрессионная модель. Система нормальных уравнений. Основы регрессионного анализа (проверка значимости коэффициентов, коэффициент детерминации Пирсона, доверительные интервалы, анализ остатков).	УК-1
4) Проблема переобучения при решении задачи восстановления регрессии. Методы борьбы с переобучением: сокращение числа параметров, регуляризация (ридж-регрессия), метод лассо. Трудоемкость методов.	УК-1
5) Метод ближайших соседей для решения задачи классификации. Теорема об оценке риска в методе ближайшего соседа.	УК-1
6) Наивный байесовский классификатор.	УК-1
7) Линейный дискриминантный анализ. Квадратичный дискриминантный анализ. Логистическая регрессия.	ОПК-2
8) Нейронные сети. Персептрон Розенблатта. Алгоритм обучения персептрона как метод стохастического градиентного спуска. Нейронные сети для решения задач классификации и восстановления регрессии. Обучение сети. Регуляризация как метод борьбы с переобучением. Понятие о глубоких нейронных сетях.	ОПК-2
9) Машина опорных векторов. Ядра и спрямляющие пространства.	ОПК-2
10) Деревья решений. Метод CART (classification and regression trees) для решения задач классификации и восстановления регрессии. Отсечения ветвей и выбор финального дерева. Методы обработки пропущенных значений.	ОПК-2
11) Ансамбли решающих правил (классификаторов). Простое и взвешенное голосование. Бустинг. Алгоритм AdaBoost. Оценка ошибки предсказания. Бустинг и аддитивные модели. Градиентный бустинг. Алгоритм градиентного бустинга деревьев решений (MART). Баггинг. Алгоритм случайных деревьев (.случайный лес.).	ОПК-2
12) Обучение без учителя. Кластеризация. Кластеризация методами теории графов. Метод центров тяжести. Метод медиан. Метод нечетких множеств. ЕМ-алгоритм.	ОПК-2

13) Иерархическая кластеризация. Агломеративные и разделяющие методы.	ОПК-2
14) Основы теории Вапника–Червоненкиса. Лемма Бернштейна. Теорема о равномерной сходимости эмпирического риска к ожидаемому риску в случае конечного класса решающих правил. Обоснование принципа минимизации эмпирического риска.	ОПК-2
15) Размерность Вапника–Червоненкиса. Лемма Зауэра. Теорема о равномерной сходимости эмпирического риска к ожидаемому риску в случае конечной размерности Вапника–Червоненкиса. Принцип структурной минимизации риска.	ОПК-2

5.2.2. Типовые задания/задачи для оценки сформированности компетенции УК-1

1. Дана обучающая выборка

x_1	0	1	1	0	0	1	1	2	6
x_2	3	3	1	0	1	1	2	3	1
y	0	0	0	0	1	1	1	1	1

Методом линейного дискриминантного анализа для каждого класса построить дискриминантную функцию и записать уравнение разделяющей поверхности.

2. Дана обучающая выборка (см. таблицу выше). Методом квадратичного дискриминантного анализа построить дискриминантные функции.
3. Дана обучающая выборка (см. таблицу выше). С помощью наивного байесова классификатора оценить вероятности $P(Y = 1 | x_1 = 1, x_1 = 2)$

5.2.3. Типовые задания/задачи для оценки сформированности компетенции ОПК-2

1. Дана обучающая выборка:

x	-1	0	0	1	2
y	1	-2	1	7	8

Методом наименьших квадратов построить полиномиальную модель вида $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$.

2. Дана обучающая выборка (см. выше). Методом ридж-регрессии построить полиномиальную модель вида $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$, если параметр регуляризации $\lambda = 2$.

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

а) Основная литература:

- 1) Вьюгин В.В. Математические основы машинного обучения и прогнозирования: Учебное пособие - Режим доступа: <http://znanium.com/catalog/product/958689>
- 2) Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые

извлекают знания из данных [Электронный ресурс] - Режим доступа:
<http://www.studentlibrary.ru/book/ISBN9785970602737.html>

б) Дополнительная литература

- 1) Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения - Режим доступа:
<http://znanium.com/catalog/product/1027758>

в) программное обеспечение и Интернет-ресурсы

- 1) Scikit-Learn: Machine Learning in Python. Международный открытый ресурс и программное обеспечение для машинного обучения [Электронный ресурс] - Режим доступа: www.scikit-learn.org
- 2) Welcome to Python.org. Международный открытый ресурс и программное обеспечение для машинного обучения [Электронный ресурс] - Режим доступа:
<https://www.python.org/>

7. Материально-техническое обеспечение дисциплины (модуля)

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: учебной мебелью, доской, проектором высокого разрешения.

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 ВО ННГУ с учетом рекомендаций ФГОС ВО по направлению 090404 Программная инженерия..

Автор: д.ф.-м.н, проф., кафедры АГДМ Золотых Н.Ю.

Рецензент: д.т.н., профессор кафедры МОСТ, Турлапов В.Е.

Заведующий кафедрой: д.ф.-м.н, проф., заведующий кафедрой АГДМ Золотых Н.Ю.

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики от 30 ноября 2022 года, протокол № 3.