

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 Ученого совета ННГУ

протокол № 10 от 02.12.2024 г.

Рабочая программа дисциплины

Машинное обучение

Уровень высшего образования

Магистратура

Направление подготовки / специальность

09.04.04 - Программная инженерия

Направленность образовательной программы

Технологии цифровой трансформации

Форма обучения

очная

г. Нижний Новгород

2025 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.О.08 Машинное обучение 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
УК-1: Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1: Знает методы критического анализа проблемных ситуаций УК-1.2: Умеет вырабатывать стратегию действий при возникновении критических ситуаций УК-1.3: Владеет основами системного подхода к анализу проблемных ситуаций.	УК-1.1: Знать методы критического анализа проблемных ситуаций машинного обучения УК-1.2: Уметь разрабатывать стратегию действий при решении задач машинного обучения УК-1.3: Владеть основами системного подхода для анализа задач машинного обучения	Тест	Экзамен: Контрольные вопросы
ОПК-2: Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач;	ОПК-2.1: Знает современные интеллектуальные технологии для решения профессиональных задач ОПК-2.2: Умеет обосновывать выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач ОПК-2.3: Имеет навыки разработки оригинальных программных средств, в том числе с использованием современных	ОПК-2.1: ЗНАТЬ основные алгоритмы и методы машинного обучения основы языка Python или среды вычислений R ОПК-2.2: УМЕТЬ использовать методы машинного обучения на практике, умения оценивать качество методов работать с библиотекой Scikit-Learn или средой для статистических вычислений R ОПК-2.3:	Контрольная работа	Экзамен: Контрольные вопросы

	интеллектуальных технологий, для решения профессиональных задач	Владеть навыками разработки методами машинного обучения для решения профессиональных задач		
--	---	--	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	4
Часов по учебному плану	144
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	32
- занятия семинарского типа (практические занятия / лабораторные работы)	32
- КСР	2
самостоятельная работа	42
Промежуточная аттестация	36 Экзамен

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
Постановка задач машинного обучения	34	10	10	20	14
Обучение с учителем	34	10	10	20	14
Обучение без учителя	38	12	12	24	14
Аттестация	36				
КСР	2			2	
Итого	144	32	32	66	42

Содержание разделов и тем дисциплины

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

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/>

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

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 - Тест) для оценки сформированности компетенции УК-1:

1. Какая из следующих задач является задачей обучения без учителя?
 - а. восстановление регрессии
 - б. классификация
 - с. кластеризация (+)
2. Какая из следующих задач является задачей обучения с учителем?
 - а. классификация (+)
 - б. кластеризация
 - с. восстановление функции распределения

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
зачтено	100-51%
не зачтено	0-50%

5.1.2 Типовые задания (оценочное средство - Контрольная работа) для оценки сформированности компетенции ОПК-2:

1. Дана обучающая выборка

x_1	0 2 1 1 1 2 4 4 4 6
x_2	4 4 3 5 4 1 0 2 1 1
y	0 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.

Критерии оценивания (оценочное средство - Контрольная работа)

Оценка	Критерии оценивания
превосходно	Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
отлично	Решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.
очень хорошо	Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами
хорошо	Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами
удовлетворительно	Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но

Оценка	Критерии оценивания
	не в полном объеме.
неудовлетворительно	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки
плохо	Отсутствие минимальных умений решения задач. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа

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	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 отдельным 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.	При решении стандартных задач не	Имеется минимальный набор	Продemonстрированы базовые	Продemonстрированы базовые	Продemonстрированы навыки	Продemonстрирован творческий

	Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	продемонстрированы базовые навыки. Имели место грубые ошибки	навыков для решения стандартных задач с некоторым и недочетами	навыки при решении стандартных задач с некоторым и недочетами	навыки при решении стандартных задач без ошибок и недочетов	при решении нестандартных задач без ошибок и недочетов	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 УК-1

1) Вероятностная постановка задачи обучения по прецедентам. Принцип минимизации эмпирического риска. Байесовская теория решений. Принцип максимума апостериорной вероятности. Байесов классификатор.

2) Экспериментальные методы оценки качества обучения.

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).

4) Проблема переобучения при решении задачи восстановления регрессии. Методы борьбы с переобучением: сокращение числа параметров, регуляризация (ридж-регрессия), метод лассо. Трудоемкость методов.

5) Метод ближайших соседей для решения задачи классификации. Теорема об оценке риска в методе ближайшего соседа.

6) Наивный байесовский классификатор.

5.3.2 Типовые задания (оценочное средство - Контрольные вопросы) для оценки сформированности компетенции ОПК-2

1) 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.

2) Машина опорных векторов. Ядра и спрямляющие пространства.

3) 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.

4) 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.).

5) Обучение без учителя. Кластеризация. Кластеризация методами теории графов. Метод центров тяжести. Метод медиан. Метод нечетких множеств. EM-алгоритм.

6) Иерархическая кластеризация. Агломеративные и разделяющие методы.

7) 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.

8) Размерность Вапника–Червоненкиса. Лемма Зауэра. Теорема о равномерной сходимости эмпирического риска к ожидаемому риску в случае конечной размерности Вапника–Червоненкиса. Принцип структурной минимизации риска.

Критерии оценивания (оценочное средство - Контрольные вопросы)

Оценка	Критерии оценивания
превосходно	свободное владение основным и дополнительным материалом без ошибок и погрешностей
отлично	свободное владение основным материалом с незначительными ошибками и погрешностями

Оценка	Критерии оценивания
очень хорошо	достаточное владение основным материалом с незначительными погрешностями
хорошо	владение основным материалом с рядом заметных погрешностей
удовлетворительно	владение минимальным материалом, необходимым по данному предмету, с рядом ошибок
неудовлетворительно	владение материалом недостаточно, необходима дополнительная подготовка. Работу за время семестра можно оценить как неудовлетворительную
плохо	отсутствие владения материалом. Работа за время семестра была оценена на «плохо»

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) 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)

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 09.04.04 - Программная инженерия.

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 02.12.2024, протокол № 5.