

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 27.08.2025

Рабочая программа дисциплины

Машинное обучение и анализ данных

Уровень высшего образования
Бакалавриат

Направление подготовки / специальность
02.03.02 - Фундаментальная информатика и информационные технологии

Направленность образовательной программы
Системное программирование

Форма обучения
очная

г. Нижний Новгород

2025 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.03 Машинное обучение и анализ данных 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-Ф2: Способен проектировать, разрабатывать, внедрять, развертывать и управлять моделями машинного обучения	<p>ПК-Ф2.1: Знает основные алгоритмы и методы машинного обучения</p> <p>ПК-Ф2.2: Знает основные концептуальные и теоретические модели искусственного интеллекта и машинного обучения</p> <p>ПК-Ф2.3: Умеет использовать методы машинного обучения на практике, оценивать качество методов, разрабатывать и анализировать концептуальные и теоретические модели машинного обучения</p>	<p>ПК-Ф2.1: ЗНАТЬ основные алгоритмы и методы машинного обучения основы языка Python или среды вычислений R</p> <p>ПК-Ф2.2: ЗНАТЬ основные концептуальные и теоретические модели искусственного интеллекта и машинного обучения</p> <p>ПК-Ф2.3: УМЕТЬ использовать методы машинного обучения на практике, оценивать качество методов работать с библиотекой Scikit-Learn или средой для статистических вычислений R, разрабатывать и анализировать концептуальные и теоретические модели машинного обучения</p>	Задачи	Зачёт: Контрольные вопросы
ППК-Р7: Способен применять искусственный интеллект (ИИ) для генерации и отладки программного кода (top)	<p>ППК-Р7.1: Применяет ИИ-инструменты для генерации программного кода</p> <p>ППК-Р7.2: Использует ИИ для анализа и отладки кода</p> <p>ППК-Р7.3: Оптимизирует код с помощью ИИ</p> <p>ППК-Р7.4: Оценивает</p>	<p>ППК-Р7.1: ППК-Р7.1. 3-1. Знает принципы работы современных генеративных ИИ-моделей для генерации кода</p> <p>ППК-Р7.1. 3-2. Знает ограничения и риски</p>	Тест	Зачёт: Контрольные вопросы

	<p>этические и профессиональные аспекты применения ИИ в разработке</p>	<p>использования ИИ-генерации (безопасность, качество кода, лицензирование)</p> <p>ППК-Р7.1. У-1. Умеет формулировать корректные текстовые запросы (промты) для генерации кода</p> <p>ППК-Р7.1. У-2. Умеет интегрировать ИИ-инструменты в среду разработки</p> <p>ППК-Р7.2:</p> <p>ППК-Р7.2. З-1. Знает методы ИИ-анализа кода</p> <p>ППК-Р7.2. З-2. Знает форматы и инструменты для автоматизированного тестирования с ИИ</p> <p>ППК-Р7.2. У-1. Умеет настраивать ИИ-инструменты для поиска уязвимостей</p> <p>ППК-Р7.2. У-2. Умеет интерпретировать рекомендации ИИ по исправлению кода</p> <p>ППК-Р7.3:</p> <p>ППК-Р7.3. З-1. Знает методы ИИ-оптимизации</p> <p>ППК-Р7.3. З-2. Знает критерии качества кода, применяемые ИИ-системами</p> <p>ППК-Р7.3. У-1. Умеет использовать ИИ для рефакторинга</p> <p>ППК-Р7.4:</p> <p>ППК-Р7.4. З-1. Знает этические нормы использования ИИ (конфиденциальность, плагиат кода и т.п.)</p> <p>ППК-Р7.4. З-2. Знает лицензионные ограничения сгенерированного кода</p> <p>ППК-Р7.4. У-1. Умеет проверять код на соответствие стандартам после ИИ-генерации</p> <p>ППК-Р7.4. У-2. Умеет документировать</p>		
--	--	---	--	--

		использование ИИ в разработке		
--	--	-------------------------------	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	2
Часов по учебному плану	72
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	16
- занятия семинарского типа (практические занятия / лабораторные работы)	48
- КСР	1
самостоятельная работа	7
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	
Постановка задач машинного обучения	14	4	8	12	2
Обучение с учителем	24	6	16	22	2
Обучение без учителя	25	6	16	22	3
Современные гетеративные ИИ-модели	4		4	4	
Введение в написание промта	4		4	4	
Аттестация	0				
КСР	1			1	
Итого	72	16	48	65	7

Содержание разделов и тем дисциплины

Цели и задачи дисциплины.

Сформировать у студентов теоретические и практические знания в области машинного обучения и анализа данных. Научить применять современные методы анализа данных для решения прикладных задач. Подготовить к самостоятельной работе с информацией и использованию ML-технологий в профессиональной деятельности. Сформировать навыки работы с современными программными средствами и архитектурами нейронных сетей.

Содержание разделов и тем дисциплины.

Вероятностная постановка задачи обучения по прецедентам. Принцип минимизации эмпирического риска. Байесовская теория решений. Принцип максимума апостериорной вероятности. Байесов классификатор. Экспериментальные методы оценки качества обучения. 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). Проблема переобучения при решении задачи восстановления регрессии. Методы борьбы с переобучением: сокращение числа параметров, регуляризация (ридж-регрессия), метод лассо. Трудоемкость методов. Метод ближайших соседей для решения задачи классификации. Теорема об оценке риска в методе ближайшего соседа. Наивный байесовский классификатор. Линейный дискриминантный анализ. Квадратичный дискриминантный анализ. Логистическая регрессия.

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. Машина опорных векторов. Ядра и спрямляющие пространства.

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. 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.). Обучение без учителя. Кластеризация. Кластеризация методами теории графов. Метод центров тяжести. Метод медиан. Метод нечетких множеств. EM-алгоритм. Иерархическая кластеризация. Агломеративные и разделяющие методы. 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. Размерность Вапника–Червоненкиса. Лемма Зауэра. Теорема о равномерной сходимости эмпирического риска к ожидаемому риску в случае конечной размерности Вапника–Червоненкиса. Принцип структурной минимизации риска.

Современные гетеративные ИИ-модели. Понятие генеративного искусственного интеллекта (Generative AI). Отличие от традиционного ИИ: анализ, классификация, прогнозирование против создания нового контента (текст, изображения, видео, музыка, код).

Принципы работы: обучение на больших массивах данных, выявление закономерностей, генерация уникальных объектов по запросу.

Популярные современные модели: GPT (OpenAI), Claude (Anthropic), Gemini (Google), Llama (Meta), Midjourney, DALL-E, Stable Diffusion. Российская экосистема: GigaChat, YandexGPT, Kandinsky.

Введение в написание промта. Особенности и ограничения современных генеративных моделей: высокая скорость генерации, качество, мультимодальность, «чёрный ящик», сложность контроля, этические вопросы.

Тенденции развития: мультимодальные модели, интеграция с реальным бизнесом, расширение сфер применения.

Практические 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.

На проведение практических занятий / лабораторных работ в форме практической подготовки отводится: очная форма обучения - 2 ч.

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/>

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 - Задачи) для оценки сформированности компетенции ПК-Ф2:

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

X_1	0 0 1 1 0 0 1 1 1 0
X_2	0 1 0 1 1 1 1 1 1 1
Y	0 0 0 0 0 1 1 1 1 1

С помощью наивного байесова классификатора оценить вероятности $P(Y = 1 | X_1 = 1, X_2 = 1)$

Критерии оценивания (оценочное средство - Задачи)

Оценка	Критерии оценивания
зачтено	Задача решена полностью или решена основная часть задачи, или задача решена с

Оценка	Критерии оценивания
	недочетами
не зачтено	Задача не решена или сделан первый этап решения задачи

5.1.2 Типовые задания (оценочное средство - Тест) для оценки сформированности компетенции ППК-Р7:

1. Что такое персептрон Розенблатта?

- а) Алгоритм кластеризации.
- б) Нейронная сеть с одним слоем, способная решать задачи линейной классификации.
- в) Метод регуляризации.
- г) Алгоритм градиентного бустинга.

2. Какой метод обучения используется для персептрона?

- а) Метод наименьших квадратов.
- б) Стохастический градиентный спуск.
- в) Метод опорных векторов.
- г) EM-алгоритм.

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
зачтено	Более 80% ответов верные
не зачтено	Менее 80% ответов верные

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	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено	Уровень знаний в объеме, соответствующем программе	Уровень знаний в объеме, соответствующем программе	Уровень знаний в объеме, соответствующем программе	Уровень знаний в объеме, превышающем программу подготовки.

	знаний вследствие отказа обучающегося от ответа		много негрубых ошибок	подготовки . Допущено несколько негрубых ошибок	подготовки . Допущено несколько несущественных ошибок	подготовк и. Ошибок нет.	
<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 и 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

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».

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-Ф2

1. 1. Сформулируйте вероятностную постановку задачи обучения по прецедентам.
2. 2. Опишите принцип минимизации эмпирического риска.
3. 3. Изложите основную идею байесовской теории решений.
4. 4. Дайте определение принципу максимума апостериорной вероятности.
5. 5. Охарактеризуйте байесовский классификатор.
6. 6. Перечислите экспериментальные методы оценки качества обучения.
7. 7. Опишите применение метода наименьших квадратов для восстановления регрессии.
8. 8. Выведите метод наименьших квадратов на основе метода максимального правдоподобия.
9. 9. Дайте определение линейной регрессионной модели.
10. 10. Представьте систему нормальных уравнений в линейной регрессии.

5.3.2 Типовые задания (оценочное средство - Контрольные вопросы) для оценки сформированности компетенции ППК-Р7

1. Перечислите основные этапы регрессионного анализа.
2. Опишите проблему переобучения при восстановлении регрессии.
3. Перечислите методы борьбы с переобучением.
4. Оцените трудоёмкость основных методов восстановления регрессии.
5. Изложите суть метода ближайших соседей для классификации.
6. Сформулируйте теорему об оценке риска в методе ближайшего соседа.
7. Охарактеризуйте наивный байесовский классификатор.
8. Сравните линейный и квадратичный дискриминантный анализ.
9. Опишите применение логистической регрессии в задачах классификации.
10. Перечислите основные этапы обучения нейронной сети.
11. Перечислите основные архитектуры, используемые в современных генеративных моделях, и приведите примеры их применения.
12. Сравните возможности и ограничения диффузионных моделей, трансформеров и GAN.
13. Назовите ключевые отличия российских и зарубежных генеративных ИИ-моделей.

14. Опишите основные сферы практического применения генеративного искусственного интеллекта и тенденции его развития.

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. Коротеев М. В. Практикум по машинному обучению на Python : учебное пособие по дисциплине: «Машинное обучение» для студентов, обучающихся по направлениям: 01.03.02 «Прикладная математика и информатика», 09.03.03 «Прикладная информатика», всех профилей (программы подготовки бакалавров). Ч. 2. Практикум по машинному обучению на Python. Часть 2 / Коротеев М. В.,Одинцова В. А.,Плешакова Е. С. - Москва : Финансовый университет, 2024. - 77 с. - Книга из коллекции Финансовый университет - Информатика.,

<https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=927158&idb=0>.

2. Платонов Алексей Владимирович. Машинное обучение : учебное пособие для вузов / А. В. Платонов. - 2-е изд. - Москва : Юрайт, 2025. - 89 с. - (Высшее образование). - ISBN 978-5-534-20732-3. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=923610&idb=0>.

3. Баланов А. Н. Машинное обучение и искусственный интеллект : учебное пособие для вузов / Баланов А. Н. - 2-е изд., стер. - Санкт-Петербург : Лань, 2025. - 172 с. - Книга из коллекции Лань - Информатика. - ISBN 978-5-507-52891-2., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=972124&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 02.03.02 - 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 25.06.2025, протокол № Протокол №11.