

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 г.

**Рабочая программа дисциплины**

Машинное обучение и анализ данных

---

Уровень высшего образования  
Бакалавриат

---

Направление подготовки / специальность  
01.03.02 - Прикладная математика и информатика

---

Направленность образовательной программы  
Математическое моделирование и искусственный интеллект

---

Форма обучения  
очная

---

г. Нижний Новгород

2025 год начала подготовки

## 1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.12 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-ТОП_8: Способен применять искусственный интеллект (ИИ) для генерации и отладки программного кода	<p>ПК-ТОП_8.1: Применяет ИИ-инструменты для генерации программного кода</p> <p>ПК-ТОП_8.2: Использует ИИ для анализа и отладки кода</p> <p>ПК-ТОП_8.3: Оптимизирует код с помощью ИИ</p> <p>ПК-ТОП_8.4: Оценивает этические и профессиональные аспекты применения ИИ в разработке</p>	<p>ПК-ТОП_8.1: Использовать современные системы автоматической генерации кода с применением алгоритмов машинного обучения и нейронных сетей. Создавать прототипы решений на основе AI-сервисов и платформ.</p> <p>ПК-ТОП_8.2: Применять методы глубокого обучения для выявления ошибок и потенциальных уязвимостей в исходном коде. Уметь использовать специализированные ИИ-решения для автоматизированного тестирования и оценки качества программного продукта.</p> <p>ПК-ТОП_8.3: Уметь оптимизировать вычислительные ресурсы с использованием механизмов прогнозирования нагрузки и поведения пользователей, улучшать скорость исполнения кода благодаря внедрению оптимизационных алгоритмов на основе анализа данных.</p>	Тест	<p>Зачёт: Контрольные вопросы</p>

		<p><b>ПК-ТОП_8.4:</b>  Владеть критическим осмыслением вопросов конфиденциальности и защиты персональных данных. Знать о правовых аспектах применения технологий ИИ в коммерческой среде, об ответственности за возможные последствия использования автоматических решений на основе ИИ.</p>		
<p><b>ПК-Ф2:</b> Способен проектировать, разрабатывать, внедрять, развертывать и управлять моделями машинного обучения</p>	<p><b>ПК-Ф2.1:</b> Знает основные алгоритмы и методы машинного обучения  <b>ПК-Ф2.2:</b> Знает основные концептуальные и теоретические модели искусственного интеллекта и машинного обучения  <b>ПК-Ф2.3:</b> Умеет использовать методы машинного обучения на практике, оценивать качество методов, разрабатывать и анализировать концептуальные и теоретические модели машинного обучения</p>	<p><b>ПК-Ф2.1:</b>  ЗНАТЬ основные алгоритмы и методы машинного обучения основы языка Python или среды вычислений R</p> <p><b>ПК-Ф2.2:</b>  ЗНАТЬ основные концептуальные и теоретические модели искусственного интеллекта и машинного обучения</p> <p><b>ПК-Ф2.3:</b>  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	<p>Зачёт:  Контрольные вопросы</p>

### 3. Структура и содержание дисциплины

#### 3.1 Трудоемкость дисциплины

	<b>очная</b>
<b>Общая трудоемкость, з.е.</b>	<b>2</b>
<b>Часов по учебному плану</b>	<b>72</b>
в том числе	
<b>аудиторные занятия (контактная работа):</b>	

- 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	
Ф Ф Ф	Ф Ф Ф	Ф Ф Ф	Ф Ф Ф	Ф Ф Ф	
Постановка задач машинного обучения	14	4	8	12	2
Обучение с учителем	28	6	20	26	2
Обучение без учителя	29	6	20	26	3
Аттестация	0				
КСР	1			1	
Итого	72	16	48	65	7

### Содержание разделов и тем дисциплины

Цели и задачи изучения дисциплины

Целью освоения дисциплины является приобретение:

- Знаний основных алгоритмов и методов машинного обучения; основ языка Python или среды вычислений R; основных концептуальных и теоретических моделей искусственного интеллекта и машинного обучения.

- Умений использовать методы машинного обучения на практике, умения оценивать качество методов

работать с библиотекой Scikit-Learn или средой для статистических вычислений R;

разрабатывать и анализировать концептуальные и теоретические модели машинного обучения.

Задачи

1) Изучить теоретические основы дисциплины «Машинное обучение и анализ данных».

2) Обеспечить формирование компетенций ПК-ТОП\_8, ПК-Ф2 в соответствии с требованиями образовательной программы.

3) Сформировать умения применять методы машинного обучения на практике, оценивать качество их работы, использовать библиотеку Scikit-Learn и/или среду для статистических вычислений R, а также разрабатывать и анализировать концептуальные и теоретические модели машинного обучения.

Содержание:

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.

#### **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 - Тест) для оценки сформированности компетенции ПК-ТОП\_8:**

1. Какая из следующих задач является задачей обучения без учителя?
  - a. восстановление регрессии
  - b. классификация
  - c. кластеризация (+)
2. Какая из следующих задач является задачей обучения с учителем?
  - a. классификация (+)
  - b. кластеризация
  - c. восстановление функции распределения
3. Переобучением называется
  - a. явление, возникающее при решении задач обучения с учителем, при котором ошибка решающего правила на тестовой выборке много больше ошибки на обучающей выборке (+)
  - b. явление, возникающее при решении задач обучения с учителем, при котором ошибка решающего правила на тестовой выборке много меньше ошибки на обучающей выборке
  - c. явление, возникающее при решении задач обучения с учителем, при котором решающее правило допускает большую ошибку на тестовой выборке

#### Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
зачтено	Набрано 100-51% верных ответов
не зачтено	Набрано 0-50% верных ответов

#### 5.1.2 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 недочетами
не зачтено	Задача не решена или сделан первый этап решения задачи

### 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	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	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несуществе	Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет.	Уровень знаний в объеме, превышающем программу подготовки.

	ответа			ошибок	нных ошибок		
<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	<b>превосходно</b>	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
	<b>отлично</b>	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично».
	<b>очень хорошо</b>	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо»
	<b>хорошо</b>	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо».
	<b>удовлетворительно</b>	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	<b>неудовлетворительно</b>	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно».
	<b>плохо</b>	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-ТОП\_8**

1. Нейронные сети. Перцептрон Розенблатта. Алгоритм обучения перцептрона 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 голосования. 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.

#### **5.3.2 Типовые задания (оценочное средство - Контрольные вопросы) для оценки сформированности компетенции ПК-Ф2**

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.

7. Линейный дискриминантный анализ. Квадратичный дискриминантный анализ. Логистическая регрессия.

### 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. Электронные ресурсы <http://www.uic.unn.ru/~zny/ml/>

Используемое лицензионное и свободно распространяемое программное обеспечение:

1. The R Project for Statistical Computing - <https://www.r-project.org>

2. Welcome to Python.org - <https://www.python.org/>

3. scikit-learn: machine learning in Python - <https://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 01.03.02 - Прикладная математика и информатика.

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.