

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО
Заместитель директора

Ю.О. Соболев

	Рабочая программа дисциплины (модуля)
по дисциплине:	Основы машинного обучения
по направлению:	Прикладная математика и информатика
профиль подготовки:	Науки о данных
	центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
	центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
курс:	1
квалификация:	магистр

Семестр, формы промежуточной аттестации: 2 (весенний) - Экзамен

Аудиторных часов: 28 всего, в том числе:

лекции: 2 час.

семинары: 26 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 167 час.

Подготовка к экзамену: 30 час.

Всего часов: 225, всего зач. ед.: 5

Программу составили:

А.В. Глинский, преподаватель

К.А. Лапин, старший методист

Программа обсуждена на заседании центра дополнительного, дополнительного профессионального и
онлайн-образования "Пуск" 01.03.2025

Аннотация

На курсе обучающиеся освоят основные навыки по разработке математических моделей, методов и алгоритмов анализа данных для машинного обучения. Курс ориентирован на формирование системных знаний и компетенций, необходимых для профессиональной деятельности и проведения научно-исследовательской работы для решения прикладных задач.

1. Цели и задачи

Цель дисциплины

- сформировать у студентов теоретическую базу в области анализа данных и машинного обучения.

Задачи дисциплины

- изучить методы анализа данных с помощью математики;
- изучить основные задачи анализа данных и машинного обучения.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1 Умеет решать задачи собственного личностного и профессионального развития, определять и реализовывать приоритеты совершенствования собственной деятельности
	УК-6.2 Оценивает свою деятельность, соотносит цели, способы и средства выполнения деятельности с её результатами
ОПК-1 Способен решать актуальные задачи фундаментальной и прикладной математики	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные научные знания и новые научные принципы и методы исследований в области прикладной математики и информатики
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
	ОПК-1.3 Понимает междисциплинарные связи в области прикладной математики и информатики и способен их применять при решении задач профессиональной деятельности
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области	ОПК-3.1 Умеет анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-3.3 Владеет аналитическими и вычислительными методами решения, задач, понимает и учитывает на практике границы применимости получаемых решений

анализ при решении задач в области профессиональной деятельности	ОПК-3.4 Способен самостоятельно приобретать, развивать и применять математические знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК-6 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-6.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-6.3 Способен использовать исследовательские методы при решении новых задач, применяя знания из различных областей науки (техники)
	ОПК-6.4 Владеет аналитическими и вычислительными методами решения, понимает и учитывает на практике границы применимости получаемых решений
ПК-8 Способен приобретать и использовать организационно-управленческие навыки в конкретной профессиональной деятельности; разрабатывать, реализовывать и управлять процессами жизненного цикла программных продуктов	ПК-8.1 Знает основы разработки и реализации процессов жизненного цикла программного обеспечения
	ПК-8.2 Умеет приобретать и использовать организационно-управленческие навыки в профессиональной деятельности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- о современном состоянии исследований в рамках тематической области своей профессиональной деятельности;
- теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты;
- принципы технико-экономического обоснования инновационных проектов.

уметь:

- анализировать проблемную ситуацию как систему, выявляя ее составляющие и связи между ними;
- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- формулировать в рамках обозначенной проблемы цель, задачи, актуальность, значимость (научную, практическую, методическую и иную в зависимости от типа проекта), ожидаемые результаты и возможные сферы их применения;
- прогнозировать результат деятельности и планировать последовательность шагов для достижения данного результата;
- представлять публично результаты проекта (или отдельных его этапов) в форме отчетов, статей, выступлений на научно-практических конференциях, семинарах и т.п.;
- оценивать актуальность исследований в области своей профессиональной деятельности и их практическую значимость;
- применять знания и навыки по использованию информационно-коммуникационных технологий для поиска и изучения научной литературы, применения прикладных программных продуктов;
- применять знания в области физико-математических наук для решения поставленной задачи, формулирования выводов и оценки полученных результатов;
- аргументированно выбирать способ проведения научного исследования;
- находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности;
- выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценить качество разработанной модели;
- разрабатывать технологические проекты, нацеленные на реализацию новых наукоемких продуктов, владеет методами планирования проектов.

владеть:

- профессиональной терминологией, используемой в современной научно-технической литературе, навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации;
- методами информационно-аналитической работы для выявления новых потребностей с целью определения наукоемких продуктов, обеспечивающих удовлетворение этих потребностей.

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

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Основные термины в машинном обучении. Базовая математика в машинном обучении	2	2		24
2	Процесс разработки проекта машинного обучения		4		24
3	Градиентный спуск для обучения дифференцируемых моделей машинного обучения		4		24
4	Линейная классификация		4		24
5	Деревья решений		4		24
6	Градиентный бустинг		4		24
7	Обучение без учителя. Кластеризация. Снижение размерности данных		4		23
Итого часов		2	26		167
Подготовка к экзамену		30 час.			
Общая трудоёмкость		225 час., 5 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 2 (Весенний)

1. Основные термины в машинном обучении. Базовая математика в машинном обучении

Общие термины. Основные компоненты. Основные типы задач. Определение вектора. Операции над векторами. Реализация в Python. Определение матрицы. Операции над матрицами. Ранг матрицы. Определитель матрицы. Реализация в Python. Определение функции. Свойства функций. Графики часто используемых функций. Обратные функции. Реализация в Python.

Определение производной. Недифференцируемые функции. Правила дифференцирования. Производные часто используемых функций. Реализация в Python.

2. Процесс разработки проекта машинного обучения

Процесс разработки проекта машинного обучения. Подготовка данных. Оценка качества алгоритмов обучения с учителем. Метод ближайших соседей.

3. Градиентный спуск для обучения дифференцируемых моделей машинного обучения

Градиентный спуск. Регуляризация. Линейная регрессия

4. Линейная классификация

Линейная классификация. Логистическая регрессия. Метод опорных векторов.

5. Деревья решений

Деревья решений. Композиции деревьев. Случайный лес.

6. Градиентный бустинг

Бустинг. XGBoost. LightGBM. Catboost.

7. Обучение без учителя. Кластеризация. Снижение размерности данных

Обучение без учителя. K-means. DBSCAN. Агломеративная кластеризация. Метод главных компонент.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

занятия проходят с применением дистанционных образовательных технологий.

6.Перечень рекомендуемой литературы

Основная литература

1. Динамическое программирование в примерах и задачах [Текст] : учеб. пособие: доп. М-вом высш. и сред. спец. образования СССР / И. Л. Калихман, М. А. Войтенко .— М : Высшая школа, 1979 .— 125 с.
2. Машинное обучение [Текст], / Х. Бринк, Дж. Ричардс, М. Феверолф, СПб., Питер, 2017

Дополнительная литература

1. Оптимизация и выпуклый анализ [Текст] = Optimisation / Ж.-Б. Ириарт-Уррути; пер. с фр. и предисл. Т. А. Бардадым под ред. А. А. Кукуш - КиевКИТ, 2004

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

Студенту для занятий потребуются:

1. Google Drive для доступа к материалам курса
2. Zoom
3. Ноутбук для участия в интерактивных занятиях.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студенту рекомендуется внимательно слушать лектора, следить за тем, что написано на доске или представлено на слайдах презентации, анализировать получаемую им информацию. В случае, если материал лекции непонятен, следует задать вопрос в отведенное для вопросов время. Студенту также рекомендуется конспектировать материал лекции в тетради, что улучшает запоминание.

При выполнении практических работ студенту рекомендуется внимательно анализировать поставленную задачу, уделяя особое внимание критериям оценки точности решения задачи. Особенное внимание следует уделять методологическим аспектам решения задач.

При ведении самостоятельной работы студенту рекомендуется внимательно подходить к изучению научных статей, обращать внимание на значимость полученного результата, на требования к обучающей выборке, на скорость работы предлагаемых алгоритмов, на результаты их сравнения с существующими. В случае, если изучаемый материал понятен не до конца, рекомендуется обращение к дополнительной литературе.

Студенту рекомендуется внимательно анализировать вопросы в итоговом тестировании.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладная математика и информатика
профиль подготовки:	Науки о данных центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
курс:	<u>1</u>
квалификация:	магистр

Семестр, формы промежуточной аттестации: 2 (весенний) - Экзамен

Разработчики:

А.В. Глинский, преподаватель
К.А. Лапин, старший методист

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1 Умеет решать задачи собственного личностного и профессионального развития, определять и реализовывать приоритеты совершенствования собственной деятельности
	УК-6.2 Оценивает свою деятельность, соотносит цели, способы и средства выполнения деятельности с её результатами
ОПК-1 Способен решать актуальные задачи фундаментальной и прикладной математики	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные научные знания и новые научные принципы и методы исследований в области прикладной математики и информатики
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
	ОПК-1.3 Понимает междисциплинарные связи в области прикладной математики и информатики и способен их применять при решении задач профессиональной деятельности
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.1 Умеет анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-3.3 Владеет аналитическими и вычислительными методами решения, задач, понимает и учитывает на практике границы применимости получаемых решений
	ОПК-3.4 Способен самостоятельно приобретать, развивать и применять математические знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК-6 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-6.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-6.3 Способен использовать исследовательские методы при решении новых задач, применяя знания из различных областей науки (техники)
	ОПК-6.4 Владеет аналитическими и вычислительными методами решения, понимает и учитывает на практике границы применимости получаемых решений
ПК-8 Способен приобретать и использовать организационно-управленческие навыки в конкретной профессиональной деятельности; разрабатывать, реализовывать и управлять процессами жизненного цикла программных продуктов	ПК-8.1 Знает основы разработки и реализации процессов жизненного цикла программного обеспечения
	ПК-8.2 Умеет приобретать и использовать организационно-управленческие навыки в профессиональной деятельности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Основы машинного обучения» обучающийся должен:

знать:

- стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- о современном состоянии исследований в рамках тематической области своей профессиональной деятельности;
- теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты;
- принципы технико-экономического обоснования инновационных проектов.

уметь:

- анализировать проблемную ситуацию как систему, выявляя ее составляющие и связи между ними;
- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- формулировать в рамках обозначенной проблемы цель, задачи, актуальность, значимость (научную, практическую, методическую и иную в зависимости от типа проекта), ожидаемые результаты и возможные сферы их применения;
- прогнозировать результат деятельности и планировать последовательность шагов для достижения данного результата;
- представлять публично результаты проекта (или отдельных его этапов) в форме отчетов, статей, выступлений на научно-практических конференциях, семинарах и т.п.;
- оценивать актуальность исследований в области своей профессиональной деятельности и их практическую значимость;
- применять знания и навыки по использованию информационно-коммуникационных технологий для поиска и изучения научной литературы, применения прикладных программных продуктов;
- применять знания в области физико-математических наук для решения поставленной задачи, формулирования выводов и оценки полученных результатов;
- аргументированно выбирать способ проведения научного исследования;
- находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности;
- выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценить качество разработанной модели;
- разрабатывать технологические проекты, нацеленные на реализацию новых наукоемких продуктов, владеет методами планирования проектов.

владеть:

- профессиональной терминологией, используемой в современной научно-технической литературе, навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации;
- методами информационно-аналитической работы для выявления новых потребностей с целью определения наукоемких продуктов, обеспечивающих удовлетворение этих потребностей.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Перечень примерных вопросов для текущего контроля:

1. Линейные уравнения.
2. Нелинейные уравнения.
3. Матрица.
4. Собственные векторы.
5. Регрессионный анализ.
6. Метод наименьших квадратов и максимальная вероятность.
7. Математическая статистика.
8. Теория оптимизации
9. Программное обеспечение для решения задач оптимизации алгоритмов на Python.

10. Библиотеки общего назначения.
11. Модели классификации.
12. Концепции теории вероятности.
13. Применение концепция теории вероятности для машинного обучения.
14. Определение регрессионного анализа.
15. Определение линейной регрессии.
16. Нелинейная регрессия.
17. Метод наименьших квадратов и максимальная вероятность.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

Примеры вопросов для промежуточной аттестации в формате экзамена:

1. Определение вектора и его основные свойства
2. Примеры задач, в которых применяются векторы в машинном обучении
3. Определение матрицы и её важные характеристики: ранг и определитель
4. Реализация операций над матрицами на языке Python с примерами
5. Определение функции и её свойства
6. Графики часто используемых функций и их обратных функций, примеры реализации на Python
7. Понятие производной и её значение в машинном обучении
8. Правила дифференцирования и производные часто используемых функций, примеры реализации на Python
9. Процесс разработки проекта машинного обучения: этапы и особенности
10. Подготовка данных: методы предобработки, очистки и преобразования данных
11. Оценка качества алгоритмов обучения с учителем: метрики и методы
12. Метод ближайших соседей: принцип работы, особенности применения и примеры
13. Градиентный спуск: основные идеи, виды и вариации
14. Регуляризация: цель, виды и примеры применения в машинном обучении

Примеры экзаменационных билетов:

Билет 1

Линейная регрессия: суть метода, принцип работы, примеры использования

Линейная классификация: основные принципы, примеры алгоритмов

Билет 2

Деревья решений: структура, принцип работы, особенности применения

Композиции деревьев: принципы построения, примеры алгоритмов (например, случайный лес)

Билет 3

Бустинг: суть метода, принцип работы, примеры алгоритмов (например, XGBoost, LightGBM)

Обучение без учителя: методы кластеризации и уменьшения размерности данных, их особенности и примеры (например, K-means, PCA)

Критерии оценивания

- оценка «отлично (10)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений
- оценка «отлично (9)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений

- оценка «отлично (8)» выставляется студенту, показавшему всесторонние систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, и правильное обоснование принятых решений
- оценка «хорошо (7)» выставляется студенту, если он твердо знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (6)» выставляется студенту, если он знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (5)» выставляется студенту, если он знает материал, и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «удовлетворительно (4)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «удовлетворительно (3)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет фрагментарно основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «неудовлетворительно (2)» выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач
- оценка «неудовлетворительно (1)» выставляется студенту, который не знает формулировок основных понятий дисциплины.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Время проведения экзамена составляет 2 академических часа.

Во время проведения экзамена обучающиеся могут пользоваться программой дисциплины.