

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

УТВЕРЖДЕНО

**Директор физтех-школы
радиотехники и компьютерных
технологий**

Д.А. Гаврилов

	Рабочая программа дисциплины (модуля)
по дисциплине:	Методы искусственного интеллекта
по направлению:	Инфокоммуникационные технологии и системы связи
профиль подготовки:	Телекоммуникационные сети и системы Физтех-школа Радиотехники и Компьютерных Технологий кафедра инфокоммуникационных систем и сетей
курс:	1
квалификация:	магистр

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

Аудиторных часов: 30 всего, в том числе:

лекции: 30 час.

семинары: 0 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 30 час.

Подготовка к экзамену: 30 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: М.В. Лебедев, канд. физ.-мат. наук, доцент

Программа обсуждена на заседании кафедры инфокоммуникационных систем и сетей 27.03.2023

Аннотация

Данный курс предназначен для получения студентами знаний к научно-исследовательской деятельности, практических навыков и компетенции в области машинного обучения. В курсе рассматриваются основные задачи обучения по прецедентам: классификация, кластеризация, регрессия, понижение размерности. Изучаются методы их решения, как классические, так и новые, созданные за последние годы.

1. Цели и задачи

Цель дисциплины

Целью курса является изучение основ теории машинного обучения и ознакомление с основными задачами обучения по прецедентам: классификация, кластеризация, регрессия, понижение размерности.

Задачи дисциплины

- Освоение основных идей и эвристик машинного обучения;
- приобретение знаний для формализации и математического описания основных проблем машинного обучения;
- анализ достоинств, недостатков и границ применимости основных методов дисциплины;
- изучение взаимосвязи между методами;
- знакомство студентов с примерами прикладных задач.

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-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
УК-4 Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	УК-4.4 Способен использовать современные средства информационно-коммуникационных технологий для академического и профессионального взаимодействия
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 представлять современную научную картину мира, выявлять естественнонаучную сущность проблем своей профессиональной деятельности, определять пути их решения и оценивать эффективность сделанного выбора	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные и прикладные научные знания в области естественных наук
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
ОПК-2 Способен реализовывать новые принципы и методы исследования современных инфокоммуникационных систем и сетей различных типов передачи, распределения, обработки и хранения информации	ОПК-2.2 Владеет навыками реализации новых принципов и методов исследования в современных инфокоммуникационных системах и сетях

ОПК-3 Способен приобретать, обрабатывать и использовать новую информацию в своей предметной области, предлагать новые идеи и подходы к решению задач своей профессиональной деятельности	ОПК-3.1 Умеет использовать современные информационные и компьютерные технологии, средства коммуникаций при поиске научно-технической информации в своей профессиональной деятельности
--	---

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- Основные понятия машинного обучения;
- основные математические модели для описания данных;
- основные задачи, а именно, регрессия, классификация, кластеризация, понижение размерности, отбор признаков.

уметь:

- Выбирать оптимальную математическую модель для описания данных;
- строить по этим моделям прогнозы и оценки для целевых признаков.

владеть:

- Методами описания моделей машинного обучения и алгоритмами решения основных задач для этих моделей.

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	Постановка задачи кластеризации. ЕМ алгоритм. K-means.	2			2
2	Иерархическая кластеризация.	2			2
3	Обнаружение аномалий.	2			2
4	Многомерное шкалирование.	2			2
5	Метод главных компонент и матричные разложения.	2			2
6	Задача регрессии.	2			2
7	Задача классификации.	2			2
8	Кросс-валидация и подбор гиперпараметров.	2			2
9	Использование ядер в машинном обучении.	2			2
10	Деревья решений.	2			2
11	Композиция алгоритмов.	2			2
12	Многоклассовая классификация.	2			2
13	Байесовские методы машинного обучения.	2			2
14	Нейронные сети.	2			2
15	Методы основанные на плотности точек.	2			2
Итого часов		30			30
Подготовка к экзамену		30 час.			

Общая трудоёмкость	90 час., 2 зач.ед.
--------------------	--------------------

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

Семестр: 2 (Весенний)

1. Постановка задачи кластеризации. EM алгоритм. K-means.

Генеративные и дискриминативные модели. EM алгоритм. Смесь гауссовских распределений
Генеративные и дискриминативные модели. EM алгоритм. Смесь гауссовских распределений (GMM). K-means. Примеры практического использования.

2. Иерархическая кластеризация.

Постановка задачи. Понятие дендрограммы.

3. Обнаружение аномалий.

Вероятностные и геометрические методы. Примеры практического использования.

4. Многомерное шкалирование.

Методы SNE и t-SNE.

5. Метод главных компонент и матричные разложения.

Сингулярное разложение матриц. Факторизация неотрицательных матриц (NMF). Примеры практического использования.

6. Задача регрессии.

Постановка задачи регрессии. Линейная регрессия. Lasso. Ridge. Использование Lasso для решения задачи compressive sensing. Примеры практического использования.

7. Задача классификации.

Постановка задачи. Логистическая регрессия. Метод опорных векторов. K-ближайших соседей. Примеры практического использования.

8. Кросс-валидация и подбор гиперпараметров.

Основные метрики оценки качества моделей. Основные виды кросс-валидации. Проверка качества моделей на отложенных выборках.

9. Использование ядер в машинном обучении.

Использования ядер в методе главных компонент и методе опорных векторов.

10. Деревья решений.

Методы основанные на деревьях.

11. Композиция алгоритмов.

Бэггинг. Случайные леса. Бустинг. Стекинг.

12. Многоклассовая классификация.

Сведение к бинарной классификации. Использование софтмакса.

13. Байесовские методы машинного обучения.

Наивный байесовский классификатор.

14. Нейронные сети.

Сверточные нейронные сети. Рекуррентные нейронные сети. Примеры практического использования.

15. Методы основанные на плотности точек.

Постановка задачи для DB-scan.

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

Учебная аудитория, оснащенная мультимедиапроектором и аудиторной доской.

6.Перечень рекомендуемой литературы

Основная литература

- 1.Бишоп Кристофер М. Распознавание образов и машинное обучение. Москва : Вильямс, 2020. 960 с.
- 2.Тибришани Роберт, Фридман Джером. Основы статистического обучения: интеллектуальный анализ данных, логический вывод и прогнозирование. Москва: Вильямс, 2020. 768 с.
- 3.Николенко С.И., Архангельская Е. О., Кадуринов А. А. Глубокое обучение. Санкт-Петербург: Питер, 2022. 480 с.

Дополнительная литература

1. Мерков А.Б. Распознавание образов. Введение в методы статистического обучения. Москва: Ленанд. 2021. 256 с.
2. Бурков А. Инженерия машинного обучения. Москва:ДМК Пресс, 2022. 306 с.

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

https://academy.yandex.ru/handbook/ml?utm_source=vk&utm_medium=cpc&utm_campaign=handbook&utm_content=ml

http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_%28%D0%BA%D1%83%D1%80%D1%81_%D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D0%B9%2C_%D0%9A.%D0%92.%D0%92%D0%BE%D1%80%D0%BE%D0%BD%D1%86%D0%BE%D0%B2%29#.D0.9A.D1.80.D0.B8.D1.82.D0.B5.D1.80.D0.B8.D0.B8_.D0.B2.D1.8B_.D0.B1.D0.BE.D1.80.D0.B0_.D0.BC.D0.BE.D0.B4.D0.B5.D0.BB.D0.B5.D0.B9_.D0.B8_.D0.BC.D0.B5.D1.82.D0.BE.D0.B4.D1.8B_.D0.BE.D1.82.D0.B1.D0.BE.D1.80.D0.B0_.D0.BF.D1.80.D0.B8.D0.B7.D0.BD.D0.B0.D0.BA.D0.BE.D0.B2

<https://alexanderdyakonov.wordpress.com/>

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

Jupyter (google colab), Anaconda (Python 3), PyTorch.

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

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

В результате изучения дисциплины студент должен знать основные определения, понятия.

Успешное освоение курса требует напряжённой самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя:

- чтение и конспектирование рекомендованной литературы;
- проработку учебного материала (по конспектам лекций, учебной и научной литературе);
- подготовку к экзамену.

Руководство и контроль за самостоятельной работой студента осуществляется в форме индивидуальных консультаций.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к преподавателю.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Инфокоммуникационные технологии и системы связи
профиль подготовки:	Телекоммуникационные сети и системы Физтех-школа Радиотехники и Компьютерных Технологий кафедра инфокоммуникационных систем и сетей
курс:	<u>1</u>
квалификация:	магистр

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

Разработчик: М.В. Лебедев, канд. физ.-мат. наук, доцент

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-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
УК-4 Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	УК-4.4 Способен использовать современные средства информационно-коммуникационных технологий для академического и профессионального взаимодействия
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 представлять современную научную картину мира, выявлять естественнонаучную сущность проблем своей профессиональной деятельности, определять пути их решения и оценивать эффективность сделанного выбора	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные и прикладные научные знания в области естественных наук
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
ОПК-2 Способен реализовывать новые принципы и методы исследования современных инфокоммуникационных систем и сетей различных типов передачи, распределения, обработки и хранения информации	ОПК-2.2 Владеет навыками реализации новых принципов и методов исследования в современных инфокоммуникационных системах и сетях
ОПК-3 Способен приобретать, обрабатывать и использовать новую информацию в своей предметной области, предлагать новые идеи и подходы к решению задач своей профессиональной деятельности	ОПК-3.1 Умеет использовать современные информационные и компьютерные технологии, средства коммуникаций при поиске научно-технической информации в своей профессиональной деятельности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Методы искусственного интеллекта» обучающийся должен:

знать:

- Основные понятия машинного обучения;
- основные математические модели для описания данных;
- основные задачи, а именно, регрессия, классификация, кластеризация, понижение размерности, отбор признаков.

уметь:

- Выбирать оптимальную математическую модель для описания данных;
- строить по этим моделям прогнозы и оценки для целевых признаков.

владеть:

- Методами описания моделей машинного обучения и алгоритмами решения основных задач для этих моделей.

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

С целью контроля освоения обучающимися учебного материала проводится устный опрос в начале занятия по теме прошлой лекции или в конце занятия по пройденной теме.

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

1. Постановка задачи кластеризации. EM алгоритм. K-means.
2. Иерархическая кластеризация. Методы, основанные на плотности точек (DB-scan).
3. Обнаружение аномалий.
4. Многомерное шкалирование. Метод SNE, t-SNE.
5. Метод главных компонент. Матричное разложение (SVD, ALS).
6. Постановка задачи регрессии. Линейная регрессия. Lasso. Ridge. Compressive sensing.
7. Кросс-валидация и подбор гиперпараметров.
8. Постановка задачи классификации. Логистическая регрессия.
9. Метод опорных векторов. K-ближайших соседей.
10. Использование ядер в машинном обучении.
11. Деревья решений.
12. Композиция алгоритмов. Бэггинг. Случайные леса. Бустинг. Стекинг.
13. Многоклассовая классификация.
14. Байесовские методы машинного обучения. Наивный Байесовский классификатор.
15. Нейронные сети. Сверточные нейронные сети. Рекуррентные нейронные сети.

Примеры билетов:

Билет №1

1. Деревья решений.
2. Использование ядер в машинном обучении.
3. Многоклассовая классификация.

Билет №2

1. Многомерное шкалирование. Метод SNE, t-SNE.
2. Иерархическая кластеризация. Методы, основанные на плотности точек (DB-scan).
3. Постановка задачи кластеризации. EM алгоритм. K-means.

Критерии оценивания

- 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;
- отлично (9) - выставляется студенту, показавшему свободное оперирование знаниями учебной программы дисциплины, выполнение заданий творческого характера;
- отлично (8) - выставляется студенту, показавшему владение программным учебным материалом с наличием несущественных ошибок в действиях, самостоятельно исправляемых учащимся;
- хорошо (7) - выставляется студенту, если он твердо знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускается в ответе или в решении задач некоторые неточности;
- хорошо (6) - выставляется студенту если он осознает воспроизведение программного учебного материала, в том числе и различной степени сложности, с несущественными ошибками, затруднения в применении отдельных навыков;

- хорошо (5) - выставляется студенту если теоретическое содержание освоено не полностью, некоторые практические навыки сформированы недостаточно, в некоторых случаях были допущены ошибки;

- 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;

- удовлетворительно (3) - выставляется студенту в случае большого количества недочетов и неправильных ответов, а также пассивной работе в ходе занятий, многие учебные задания не выполнены;

- 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;

- неудовлетворительно (1) - выставляется студенту, который не освоил теоретическое и практическое содержание курса, все выполненные учебные задания содержат грубые ошибки.

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

При проведении устного экзамена обучающемуся предоставляется 60 на подготовку. Опрос обучающегося по билету на устном экзамене не должен превышать двух астрономических часов. Во время проведения зачета обучающиеся могут пользоваться программой дисциплины.