Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский ядерный университет «МИФИ»

ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ

КАФЕДРА КИБЕРНЕТИКИ

ОДОБРЕНО УМС ИИКС

Протокол № 8/1/2025

от 25.08.2025 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

ВВЕДЕНИЕ В МАШИННОЕ ОБУЧЕНИЕ

Направление подготовки (специальность) [1] 09.03.04 Программная инженерия

Семестр	**Трудоемкость, кред.**	**Общий объем курса, час.**	**Лекции, час.**	**Практич. занятия, час.**	**Лаборат. работы, час.**	**В форме практической подготовки, час.**	**СРС, час.**	**КСР, час.**	**Форма(ы) контроля, экз./зач./КР/КП**
6	2-4	72-144	30	15	15		12-48	0	Э , ЗО
Итого	2-4	72-144	30	15	15	0	12-48	0	

АННОТАЦИЯ

Машинное обучение – раздел искусственного интеллекта, изучающий методы построения моделей и алгоритмов, способных обучаться. Методы машинного обучения используются при решении широкого круга прикладных задач, для которых разработка явного алгоритма решения затруднительна или невозможна. И этот круг задач постоянно расширяется. Повсеместная информатизация приводит к накоплению огромных объёмов данных в науке, производстве, бизнесе, транспорте, здравоохранении.

В настоящем курсе рассматриваются основные принципы машинного обучения с учителем, изучаются постановки задач регрессии и классификации и используемые для их решения модели и методы обучения, а также оценки качества обученных моделей, рассматриваются особенности организации процесса обучения с учителем и применения методов машинного обучения для решения практических задач. Изложение ведется строгим математическим языком, сопровождается множеством формул и математических выкладок.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Целью учебного курса является ознакомление студентов с основными принципами и современными математическими и алгоритмическими подходами к построению моделей машинного обучения.

Учебный курс ставит также своей целью приобретение умений ставить практические задачи в терминах машинного обучения, освоение студентами практических приемов построения моделей машинного обучения и оценки их качества. Выполняемое студентами домашнее задание позволит получить практический опыт по созданию и экспериментальному исследованию свойств обучающихся моделей и методов обучения.

2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО

Дисциплина требует от слушателя общематематической подготовки по дискретной математике, математическому анализу, линейной алгебре, методам оптимизации, а также углубленного знания теории вероятностей и математической статистики. Полученные знания могут быть применены студентами в процессе обучения при выполнении текущей научно-исследовательской и выпускной работ.

Дисциплина должна изучаться после прохождения курса «Математическая статистика». В свою очередь, дисциплина является предшествующей и необходимой для изучения следующих курсов профессионального цикла подготовки:

- Анализ данных и машинное обучение: часть 1;
- Машинное зрение.

Дисциплина формирует систему базовых понятий, необходимых для специалиста в области программной инженерии и больших данных, прикладной математики и информатики, способствует выработке профессиональных решений практических задач.

3. ФОРМИРУЕМЫЕ КОМПЕТЕНЦИИ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Универсальные и(или) общепрофессиональные компетенции:

Код и наименование компетенции	**Код и наименование индикатора достижения компетенции**

Профессиональные компетенции в соответствии с задачами и объектами (областями знаний) профессиональной деятельности:

Задача профессиональной деятельности (ЗПД)	**Объект или область знания**	**Код и наименование профессиональной компетенции; Основание (профессиональный стандарт-ПС, анализ опыта)**	**Код и наименование индикатора достижения профессиональной компетенции**
производственно-технологический			
- освоение и применение средств автоматизированного проектирования, разработки, тестирования и сопровождения программного обеспечения; - освоение и применение методов и инструментальных средств управления инженерной деятельностью и процессами жизненного цикла программного обеспечения; - использование типовых методов для контроля, оценки и обеспечения качества программной продукции; - обеспечение соответствия разрабатываемого программного обеспечения и технической документации российским и международным стандартам, техническим условиям, ведомственным нормативным документам и стандартам предприятия; - участие	- программный продукт (создаваемое программное обеспечение) - процессы жизненного цикла программного продукта - методы и инструменты разработки программного продукта	ПК-1 [1] - способен применять основные методы и инструменты разработки программного обеспечения *Основание:* Профессиональный стандарт: 06.001, 06.017	З-ПК-1[1] - Знать основные понятия из области разработки программных систем, применяемые метрики, методы и инструментальные средства.; У-ПК-1[1] - Уметь применять основные методы разработки программного обеспечения; применять основные инструменты разработки программного обеспечения; В-ПК-1[1] - Владеть основными методами разработки программного обеспечения; основными инструментами разработки программного обеспечения

в процессах разработки программного обеспечения			
научно-исследовательский			
- участие в проведении научных исследований (экспериментов, наблюдений и количественных измерений), связанных с объектами профессиональной деятельности (программными продуктами, проектами, процессами, методами и инструментами программной инженерии), в соответствии с утвержденными заданиями и методиками; - построение моделей объектов профессиональной деятельности с использованием инструментальных средств компьютерного моделирования; - составление описания проводимых исследований, подготовка данных для составления обзоров и отчетов;	- программный проект (проект разработки программного продукта) - программный продукт (создаваемое программное обеспечение) - процессы жизненного цикла программного продукта - методы и инструменты разработки программного продукта	ПК-11 [1] - способен к формализации в своей предметной области с учетом ограничений используемых методов исследования *Основание:* Профессиональный стандарт: 40.011	З-ПК-11[1] - Знать методы формализации в своей предметной области с учетом ограничений используемых методов исследования; У-ПК-11[1] - Уметь формализовать в своей предметной области ; В-ПК-11[1] - Владеть методами формализации в своей предметной области с учетом ограничений используемых методов исследования
- участие в проведении научных исследований (экспериментов, наблюдений и количественных измерений), связанных с объектами профессиональной деятельности (программными продуктами, проектами, процессами, методами и инструментами программной инженерии), в	- программный проект (проект разработки программного продукта) - программный продукт (создаваемое программное обеспечение) - процессы жизненного цикла программного продукта -	ПК-12 [1] - способен использовать методы и инструментальные средства исследования объектов профессиональной деятельности *Основание:* Профессиональный стандарт: 40.011	З-ПК-12[1] - Знать методы исследования объектов профессиональной деятельности; инструментальные средства исследования объектов профессиональной деятельности; У-ПК-12[1] - Уметь применять методы и инструментальные средства исследования объектов

соответствии с утвержденными заданиями и методиками; - построение моделей объектов профессиональной деятельности с использованием инструментальных средств компьютерного моделирования; - составление описания проводимых исследований, подготовка данных для составления обзоров и отчетов;	методы и инструменты разработки программного продукта		профессиональной деятельности; В-ПК-12[1] - Владеть методами и инструментальными средствами исследования объектов профессиональной деятельности

4. ВОСПИТАТЕЛЬНЫЙ ПОТЕНЦИАЛ ДИСЦИПЛИНЫ

Направления/цели воспитания	**Задачи воспитания (код)**
Профессиональное воспитание	Создание условий, обеспечивающих, формирование ответственности за профессиональный выбор, профессиональное развитие и профессиональные решения (В18)
Профессиональное воспитание	Создание условий, обеспечивающих, формирование научного мировоззрения, культуры поиска нестандартных научно-технических/практических решений, критического отношения к исследованиям лженаучного толка (В19)

5. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Разделы учебной дисциплины, их объем, сроки изучения и формы контроля:

№ п.п	**Наименование раздела учебной дисциплины**	**Недели**	**Лекции/ Практ. (семинары)/ Лабораторные работы, час.**	**Обязат. текущий контроль (форма*, неделя)**	**Максимальный балл за раздел****	**Аттестация раздела (форма*, неделя)**	**Индикаторы освоения компетенции**
	6 Семестр						
1	Обучение с учителем: основные принципы	1-8	16/8/8		25	КИ-8	В-ПК-12, З-ПК-1, У-ПК-1,

							В-ПК-1, З-ПК-11, У-ПК-11, В-ПК-11, З-ПК-12, У-ПК-12
2	Обучение с учителем: классификация и регрессия	9-15	14/7/7		25	КИ-15	З-ПК-1, У-ПК-1, В-ПК-1, З-ПК-11, У-ПК-11, В-ПК-11, З-ПК-12, У-ПК-12, В-ПК-12
	Итого за 6 Семестр		30/15/15		50		
	Контрольные мероприятия за 6 Семестр				50	Э, ЗО	З-ПК-1, У-ПК-1, В-ПК-1, З-ПК-11, У-ПК-11, В-ПК-11, З-ПК-12, У-ПК-12, В-ПК-12, З-ПК-1, У-ПК-1, В-ПК-1, З-ПК-11, У-ПК-11, В-ПК-11, З-ПК-12, У-ПК-12, В-ПК-12

* – сокращенное наименование формы контроля

** – сумма максимальных баллов должна быть равна 100 за семестр, включая зачет и (или) экзамен

Сокращение наименований форм текущего контроля и аттестации разделов:

Обозначение	**Полное наименование**
КИ	Контроль по итогам
ЗО	Зачет с оценкой
З	Зачет

КАЛЕНДАРНЫЙ ПЛАН

Недели	**Темы занятий / Содержание**	**Лек., час.**	**Пр./сем., час.**	**Лаб., час.**
	6 Семестр	30	15	15

1-8	**Обучение с учителем: основные принципы**	16	8	8
1 - 2	**Основные парадигмы машинного обучения** Машинное обучение и науки о данных. Data-driven и model-based подходы. История машинного обучения. Общая схема решения задач машинного обучения. Стандарт CRISP-DM. Типы задач машинного обучения. Примеры прикладных задач. Основные понятия и определения. Обзор основных парадигм машинного обучения.	Всего аудиторных часов		
		4	2	2
		Онлайн		
		0	0	0
3 - 5	**Обучение с учителем: основные принципы** Методы обучения с учителем. Обучающийся алгоритм. Индуктивный порог. Проблема переобучения и понятие обобщающей способности. Функция потерь. Теоретический и эмпирический риск. Принцип минимизации эмпирического риска. Смещение и дисперсия ошибки обучаемой модели. Разложение «bias-variance». Связь сложности модели, индуктивного порога и способности к обобщению. Оценивание точности обученных моделей. Методы кросс-валидации: hold-out, k-fold, LOOCV. Стратификация выборки при кросс-валидации, внутренняя кросс-валидация.	Всего аудиторных часов		
		6	3	3
		Онлайн		
		0	0	0
6 - 8	**Обучение с учителем: регрессия** Понятие статистической модели. Теоретические и статистические модели. Примеры. Регрессионные модели. Постановка задачи обучения регрессионной модели. Риск модели с квадратичной функцией потерь. Функция регрессии. Оптимальность регрессионных моделей. Задачи, виды и этапы регрессионного анализа. Метод наименьших квадратов. Матрица плана и система нормальных уравнений. МНК-оценки параметров простейшей линейной регрессии. Ошибки и остатки модели. Предположения регрессионного анализа. Свойства МНК-оценок: линейность, состоятельность, несмещенность, эффективность, нормальность. Теорема Гаусса-Маркова. Метод наименьших квадратов. Проекционная матрица. МНК-оценки параметров множественной линейной регрессии. Статистические характеристики МНК-оценок. Проблема гетероскедастичности. Статистические тесты на гетероскедастичность. Визуальный анализ гетероскедастичности остатков регрессионной модели. Подходы к устранению гетероскедастичности. Взвешенный метод наименьших квадратов. Робастный регрессионный анализ. Метод наименьших модулей, LAR-регрессия. Постановка задачи и обучение. Использование взвешенного метода наименьших квадратов для обучения робастных моделей. Метод итеративного взвешивания (Iteratively Reweighted Least-Squares, IRLS).	Всего аудиторных часов		
		6	3	3
		Онлайн		
		0	0	0
9-15	**Обучение с учителем: классификация и регрессия**	14	7	7
9 - 11	**Диагностика регрессионных моделей** Цели и виды диагностики регрессионных моделей. Проверка линейности. Графики остатков и частичной	Всего аудиторных часов		
		6	3	3
		Онлайн		

	регрессии. Смещение оценок вследствие неучтенных переменных (omitted variable bias). Проверка гетероскедастичности, независимости и нормальности регрессионных остатков. Визуальная и количественная диагностика. Квантильные диаграммы. Понятие регрессионного выброса. Показатели аномальности наблюдения. Диаграмма box-and-whisker. Оценка влияния наблюдений на регрессионную модель. Связь аномальности с левериджами и стандартизованными регрессионными остатками. Расстояние Кука. Понятие мультиколлинеарности. Негативные эффекты мультиколлинеарности регрессоров. Показатели мультиколлинеарности. Variance inflation factor (VIF). Методы устранения мультиколлинеарности. Регуляризованные регрессионные модели. Гребневая регрессия. LASSO-регрессия. Подходы к выбору регрессоров. Пошаговая регрессия.	0	0	0
12 - 14	**Бинарная классификация** Постановка задачи классификации. Бинарная классификация. Функция потерь и эмпирический риск классификатора. Виды функций потерь. Статистический взгляд на задачу обучения. Принцип минимизации эмпирического риска и метод максимального правдоподобия. Регуляризованный эмпирический риск. Оценка точности бинарного классификатора. Матрица ошибок. Специфичность и чувствительность, точность и полнота. ROC-анализ и PR-анализ. Показатели ROC AUC и PR AUC.	Всего аудиторных часов		
		6	3	3
		Онлайн		
		0	0	0
15	**Многоклассовая классификация** Постановка задачи многоклассовой классификации. Подходы к многоклассовой классификации. Метод ECOC. Схемы кодирования. Методы декодирования. Оценка точности многоклассового классификатора. Матрица ошибок. Микро- и макро- показатели точности. Специфичность и чувствительность, точность и полнота. Микро- и макро- ROC-анализы и PR-анализы.	Всего аудиторных часов		
		2	1	1
		Онлайн		
		0	0	0

Сокращенные наименования онлайн опций:

Обозначение	**Полное наименование**
ЭК	Электронный курс
ПМ	Полнотекстовый материал
ПЛ	Полнотекстовые лекции
ВМ	Видео-материалы
АМ	Аудио-материалы
Прз	Презентации
Т	Тесты
ЭСМ	Электронные справочные материалы
ИС	Интерактивный сайт

ТЕМЫ ЛАБОРАТОРНЫХ РАБОТ

Недели	Темы занятий / Содержание
	6 Семестр
1 - 8	**Регрессия** Регрессия
9 - 16	**Классификация** Классификация

6. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

1. Лекционные занятия:

a. комплект электронных презентаций/слайдов,

b. аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук)

2. Практические и лабораторные занятия:

a. компьютерный класс,

b. презентационная техника (проектор, экран, компьютер/ноутбук).

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Фонд оценочных средств по дисциплине обеспечивает проверку освоения планируемых результатов обучения (компетенций и их индикаторов) посредством мероприятий текущего, рубежного и промежуточного контроля по дисциплине.

Связь между формируемыми компетенциями и формами контроля их освоения представлена в следующей таблице:

Компетенция	Индикаторы освоения	Аттестационное мероприятие (КП 1)
ПК-1	З-ПК-1	ЗО, Э, КИ-8, КИ-15
	У-ПК-1	Э, ЗО, КИ-8, КИ-15
	В-ПК-1	ЗО, Э, КИ-8, КИ-15
ПК-11	З-ПК-11	ЗО, Э, КИ-8, КИ-15
	У-ПК-11	ЗО, Э, КИ-8, КИ-15
	В-ПК-11	Э, ЗО, КИ-8, КИ-15
ПК-12	З-ПК-12	ЗО, Э, КИ-8, КИ-15
	У-ПК-12	Э, ЗО, КИ-8, КИ-15
	В-ПК-12	Э, ЗО, КИ-8, КИ-15

Шкалы оценки образовательных достижений

Шкала каждого контрольного мероприятия лежит в пределах от 0 до установленного максимального балла включительно. Итоговая аттестация по дисциплине оценивается по 100-балльной шкале и представляет собой сумму баллов, заработанных студентом при выполнении заданий в рамках текущего и промежуточного контроля.

Итоговая оценка выставляется в соответствии со следующей шкалой:

Сумма баллов	Оценка по 4-х балльной шкале	Отметка о зачете	Оценка ECTS

90-100	5 – *«отлично»*	*«зачтено»*	A
85-89	4 – «*хорошо*»		B
75-84			C
70-74			D
65-69	3 – «*удовлетворительно*»		
60-64			E
ниже 60	2 – «*неудовлетворительно*»	*«не зачтено»*	F

Оценка «отлично» соответствует глубокому и прочному освоению материала программы обучающимся, который последовательно, четко и логически стройно излагает свои ответы, умеет тесно увязывать теорию с практикой, использует в ответах материалы монографической литературы.

Оценка «хорошо» соответствует твердым знаниям материала обучающимся, который грамотно и, по существу, излагает свои ответы, не допуская существенных неточностей.

Оценка «удовлетворительно» соответствует базовому уровню освоения материала обучающимся, при котором освоен основной материал, но не усвоены его детали, в ответах присутствуют неточности, недостаточно правильные формулировки, нарушения логической последовательности.

Отметка «зачтено» соответствует, как минимум, базовому уровню освоения материала программы, при котором обучающийся владеет необходимыми знаниями, умениями и навыками, умеет применять теоретические положения для решения типовых практических задач.

Оценку «неудовлетворительно» / отметку «не зачтено» получает обучающийся, который не знает значительной части материала программы, допускает в ответах существенные ошибки, не выполнил все обязательные задания, предусмотренные программой. Как правило, такие обучающиеся не могут продолжить обучение без дополнительных занятий.

8. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

ОСНОВНАЯ ЛИТЕРАТУРА:

1. ЭИ Г 77 Data Science. Наука о данных с нуля: Пер. с англ. — 2-е изд., перераб. и доп. : , Грас Д., Санкт-Петербург: БХВ-Петербург, 2021

2. ЭИ П 37 Python для сложных задач: наука о данных и машинное обучение : , Плас Дж. Вандер, Санкт-Петербург: Питер, 2018

3. ЭИ М 63 Введение в анализ данных : учебник и практикум, Миркин Б. Г., Москва: Юрайт, 2022

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА:

1. ЭИ Ф 70 Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных : , Флах П., Москва: ДМК Пресс, 2015

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:

Специальное программное обеспечение не требуется

LMS И ИНТЕРНЕТ-РЕСУРСЫ:

https://online.mephi.ru/

http://library.mephi.ru/

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Специальное материально-техническое обеспечение не требуется

10. УЧЕБНО-МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ СТУДЕНТОВ

В качестве оценочного средства используется 100 бальная семестровая система, учитывающая посещаемость занятий, активность (выполнение домашних занятий), выполнение тематических домашних заданий по каждому разделу, контрольно-тестовая работа по каждому разделу.

Итоговый балл за раздел (КИ) формируется следующим образом:
посещаемость семинарских занятий (еженедельно) не менее 80% +2 балла
не менее 50% +1 балл
менее 50% 0 баллов
ДЗ – выполнения тематического ДЗ (по каждому разделу)
Выполнено не менее 80% +3 баллов
Выполнено не менее 50% +2 балла
Выполнено менее 50% 0 баллов
Самостоятельная работа студента включает:
- Повторение теоретического материала
- Выполнение ДЗ

11. УЧЕБНО-МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРЕПОДАВАТЕЛЕЙ

В качестве оценочного средства используется 100 бальная семестровая система, учитывающая посещаемость занятий, активность (выполнение домашних занятий), выполнение тематических домашних заданий по каждому разделу, контрольно-тестовая работа по каждому разделу. Каждый раздел проходит аттестацию.

Итоговый балл за раздел (КИ) формируется следующим образом:
посещаемость семинарских занятий (еженедельно) не менее 80% +2 балла
не менее 50% +1 балл
менее 50% 0 баллов
ДЗ – выполнения тематического ДЗ (по каждому разделу)
Выполнено не менее 80% +3 баллов

Выполнено не менее 50% +2 балла
Выполнено менее 50% 0 баллов

Автор(ы):

Трофимов Александр Геннадьевич, к.т.н.