

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ))



Рабочая программа дисциплины (модуля),
как компонент образовательной программы
высшего образования - программы бакалавриата
по направлению подготовки
09.03.01 Информатика и вычислительная техника,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Методы оптимизации в машинном обучении

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): IT-сервисы и технологии обработки данных на транспорте

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 170737
Подписал: заместитель директора академии Паринов Денис Владимирович
Дата: 13.06.2024

1. Общие сведения о дисциплине (модуле).

Целью освоения учебной дисциплины (модуля) являются применения методов оптимизации для построения эффективных моделей машинного обучения при создании и сопровождении программного обеспечения информационно-вычислительных систем различного назначения;

Задачами освоения дисциплины являются:

- овладение основными понятиями, идеями и методами в сфере проблематики и современных технологий автоматизации машинного обучения,
- приобретение практических навыков применения методов оптимального представления и обработки знаний в этих системах, методов и алгоритмов обработки информации и оптимального управления,
- обеспечение качества в проектах в области ИТ в соответствии с установленными регламентами.

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

ОПК-3 - Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности;

ОПК-4 - Способен участвовать в разработке стандартов, норм и правил, а также технической документации, связанной с профессиональной деятельностью;

ОПК-9 - Способен осваивать методики использования программных средств для решения практических задач;

ПК-1 - Способен анализировать большие данные с использованием существующей в организации методологической и технологической инфраструктуры.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

- основные методы оптимального представления и обработки больших данных в системах искусственного интеллекта (ИИ),

- основные понятия, идеи и методы в сфере проблематики и современных технологий автоматизации машинного обучения,
- методы оптимизации для построения эффективных моделей машинного обучения при создании и сопровождении программного обеспечения,
- методы оптимизации систем в условиях неопределенности.

Уметь:

- разрабатывать и проводить сравнительный анализ алгоритмов распознавания и обработки больших данных,
- осуществлять моделирование при поиске оптимальных решений,
- разрабатывать математические модели процессов и объектов, методы их исследования, выполнять их сравнительный анализ,
- интерпретировать реальные задачи как задачи оптимизации, записывать их формальную математическую постановку

Владеть:

- навыками применения методов машинного обучения при решении профессиональных задач,
- построения оптимизационных моделей в различных предметных областях,
- математической формализации естественнонаучных проблем, исследования получаемых моделей и содержательного анализа результатов,
- прикладными программными инструментами в области машинного обучения, в том числе глубокого обучения.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 4 з.е. (144 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №5
Контактная работа при проведении учебных занятий (всего):	80	80
В том числе:		
Занятия лекционного типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 64 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Тема 1. Задачи оптимизации введение Рассматриваемые вопросы: - история появления и развития задач оптимизации - основные методы оптимизации - вычислительные мощности необходимые для решения задач оптимизации
2	Тема 2. Задачи оптимизации в машинном обучении Рассматриваемые вопросы: - оптимизационный подход в машинном обучении - разреженные методы машинного обучения
3	Тема 3. Методы внутренней точки и отсекающих плоскостей Рассматриваемые вопросы: - Методы внутренней точки - Методы отсекающих плоскостей
4	Тема 4. Стохастическая оптимизация Рассматриваемые вопросы: - Методы оптимизации с использованием глобальных верхних оценок, зависящих от параметра - Стохастическая оптимизация
5	Тема 5. Градиентный спуск. Начало Рассматриваемые вопросы: - стохастический градиентный спуск - импульс

№ п/п	Тематика лекционных занятий / краткое содержание
6	<p>Тема 6. Градиентный спуск. Модификации. Часть 1.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Метод Чебышева, - Метод сопряженных градиентов - Метод Нестерова
7	<p>Тема 7. Градиентный спуск. Модификации. Часть 2.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Субградиентный спуск - Proximal методы - SGD с импульсом и Nesterov Accelerated Gradient
8	<p>Тема 8. Градиентный спуск. Модификации. Часть 3.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Субградиентный спуск - Proximal методы - SGD с импульсом и Nesterov Accelerated Gradient
9	<p>Тема 9. Адаптивные методы оптимизации</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - параметры модели - скорость обучения - AdaGrad
10	<p>Тема 10. Оптимизация с помощью перебора и случайным поиском.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Оценка качества моделей машинного обучения - Подбор гипер параметров модели и кросс-валидация - Оптимизация гиперпараметров моделей машинного обучения методами поиска по решетке и случайного поиска
11	<p>Тема 11. Эволюционные алгоритмы.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - применение эволюционных методов для решения задачи оптимизации гипер параметров моделей - настройка гипер параметров моделей машинного обучения эволюционными алгоритмами - углубленное изучение теоретического материала по теме раздела, ответы на вопросы
12	<p>Тема 12. Мета-моделирование</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - ансамбли моделей машинного обучения - ансамбли моделей машинного обучения для настройки гипер параметров моделей машинного обучения
13	<p>Тема 13. Байесовская оптимизация</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Байесовская оптимизация - Настройка гиперпараметров моделей машинного обучения с использованием байесовской оптимизации
14	<p>Тема 14. Выбор оптимального метода машинного обучения</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Подходы к выбору эффективной модели машинного обучения при решении практических задач, - Построение эффективной модели машинного обучения для практической задачи

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	<p>Тема 1. Задачи оптимизации введение</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - примеры задач на оптимизацию - примеры задач, которые можно решить ручным расчетом - примеры задач, решаемых с применением компьютерных вычислительных мощностей
2	<p>Тема 2. Задачи оптимизации</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - программное обеспечение для решения задач оптимизации - встроенные методы оптимизации - настройка методов оптимизации
3	<p>Тема 3. Задачи оптимизации в машинном обучении</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - оптимизационный подход в машинном обучении - разреженные методы машинного обучения
4	<p>Тема 4. Разреженные данные и разреженные методы</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - разреженные данные - почему машинное обучение плохо работает с разреженными данными - устойчивость моделей машинного обучения к разреженным данным - исследования влияния разреженных данных на результаты обучения
5	<p>Тема 5. Разреженные методы машинного обучения</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - разреженные РСА - разреженные нейронные сети
6	<p>Тема 6. Методы внутренней точки и отсекающих плоскостей</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - задача с ограничениями - методы внутренней точки - задача линейного программирования - задача нелинейного программирования
7	<p>Тема 7. Методы внутренней точки и отсекающих плоскостей</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Методы отсекающих плоскостей - метод ветвей и границ - симплекс метод
8	<p>Тема 8. Стохастическая оптимизация</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Методы оптимизации с использованием глобальных верхних оценок, зависящих от параметра - примеры задач
9	<p>Тема 9. Стохастическая оптимизация</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - общий случай - Стохастическая оптимизация для обучения без учителя
10	<p>Тема 10. Градиентный спуск. Начало</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - метод горячо холодного - корректировка весов нейронной сети с несколькими хода

№ п/п	Тематика практических занятий/краткое содержание
	<ul style="list-style-type: none"> - шаг обучения - метод наискорейшего спуска
11	<p>Тема 11. Стохастический градиентный спуск</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - функция потерь - локальный минимум - стохастический градиент - отличие наискорейшего спуска от стохастического
12	<p>Тема 12. Метод импульсов</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - экспоненциальное скользящее среднее - импульс Нестерова - RMSProp
13	<p>Тема 13. Субградиентный спуск</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - задача - определение размера шага - программная реализация
14	<p>Тема 14. Proximal методы</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - задача - обобщение проективных градиентных методов - регуляризация - программная реализация
15	<p>Тема 15. Основные метода градиентного спуска</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Метод Чебышева - Метод сопряженных градиентов
16	<p>Тема 16. Адаптивные методы оптимизации</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Среднеквадратичное распространение - сочетание идеи RMSProp и оптимизатора импульса. - Адаптивная оценка момента (Adam).
17	<p>Тема 17. Оптимизация с помощью перебора</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - задача перебора вибропараметров. Сетка параметров. - кросс-валидация, программная реализация - оценка времени на перебор
18	<p>Тема 18. Случайный поиск</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - постановка задачи - отличия от перебора - кросс-валидация на случайном поиске
19	<p>Тема 19. Эволюционные алгоритмы</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - генетические алгоритмы - генетические алгоритмы на нейронных сетях - оптимизация генетического алгоритма
20	<p>Тема 20. Генетические алгоритмы, практическое применение</p> <p>Рассматриваемые вопросы:</p>

№ п/п	Тематика практических занятий/краткое содержание
	- задачи коммивояжера - библиотека PyGAD - примеры задач
21	Тема 21. Мета-моделирование Рассматриваемые вопросы: - ансамбли моделей машинного обучения - ансамбли моделей машинного обучения для настройки гипер параметров моделей машинного обучения - ансамбли catboost
22	Тема 22. Байесовская оптимизация Рассматриваемые вопросы: - Байесовская оптимизация - Настройка гиперпараметров моделей машинного обучения с использованием байесовской оптимизации
23	Тема 23. Выбор оптимального метода машинного обучения Рассматриваемые вопросы: - Подходы к выбору эффективной модели машинного обучения при решении практических задач, - Построение эффективной модели машинного обучения для практической задачи

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Работа с учебной литературой
2	Участие в онлайн-конференциях и мастер-классах
3	Поиск алгоритмов обработки данных в открытых источниках
4	Выполнение курсовой работы.
5	Подготовка к промежуточной аттестации.
6	Подготовка к текущему контролю.
7	Выполнение курсовой работы.
8	Подготовка к промежуточной аттестации.
9	Подготовка к текущему контролю.

4.4. Примерный перечень тем курсовых работ

1. Использование метода опорных векторов в задачах машинного обучения.
2. Использование метода наименьших квадратов в задачах машинного обучения.
3. Использование нейронных сетей долгой краткосрочной памяти в задачах анализа временных рядов.
4. Использование библиотек Python в задачах машинного обучения, для создания когнитивных приложений, в задачах распознавания образов.

5. Обработка естественного языка в когнитивных приложениях.
6. Алгоритмы поиска ассоциативных правил.
7. Рекомендательные системы.
8. Распознавание образов с использованием искусственных нейронных сетей глубокого обучения.
9. Обучение с подкреплением в системах искусственного интеллекта.
10. Сегментация изображений с использованием методов кластерного анализа.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль ; перевод с английского А. А. Слинкина. — 2-е изд. — Москва : ДМК Пресс, 2018. — 652 с. — ISBN 978-5-97060-618-6	https://e.lanbook.com/book/107901
2	Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7	https://e.lanbook.com/book/69955

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

- <https://habr.com/ru> - база знаний в виде статей, обзоров
- <https://journal.tinkoff.ru/short/ai-for-all/> - база данных нейронных сетей
- <https://vc.ru/services/916617-luchshie-neyroseti-bolshaya-podborka-iz-top-200-ii-generatorov-po-kategoriyam> - база данных нейронных сетей
- <https://github.com/abalmumcu/bert-rest-api> - профессиональная платформа для командой работы над проектов (нейронная сеть bert)
- <http://library.miit.ru/> - электронно-библиотечная система Научно-технической библиотеки МИИТ
- <https://proglib.io/p/raspoznavanie-obektov-s-pomoshchyu-yolo-v3-na-tensorflow-2-0-2020-11-08> - профессиональная библиотека программистов
- https://yandex.cloud/ru/blog/posts/2022/12/andrey-berger-and-yandex-cloud?utm_referrer=https%3A%2F%2Fyandex.ru%2F - библиотека профессиональных статей разработчиков Яндекс

<https://yandex.cloud/ru/blog> - библиотека профессиональных статей разработчиков Яндекс

<https://tproger.ru/translations/opencv-python-guide> - библиотека основных команд OpenCV

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

Microsoft Office
библиотеки Python

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

Компьютер преподавателя
Компьютеры студентов
экран для проектора, маркерная доска,
Проектор

9. Форма промежуточной аттестации:

Курсовая работа в 5 семестре.
Экзамен в 5 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, к.н. Академии "Высшая
инженерная школа"

Б.В. Игольников

Согласовано:

Заместитель директора академии

Д.В. Паринов

Председатель учебно-методической
комиссии

Д.В. Паринов