

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 магистратуры
по направлению подготовки
09.04.01 Информатика и вычислительная техника,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Развертывание и поддержка моделей искусственного интеллекта

Направление подготовки: 09.04.01 Информатика и вычислительная техника

Направленность (профиль): Искусственный интеллект и предиктивная аналитика в транспортных системах

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: 5665
Подписал: заведующий кафедрой Нутович Вероника
Евгеньевна
Дата: 01.09.2024

1. Общие сведения о дисциплине (модуле).

Целью освоения данной дисциплины является получение базовых, теоретических знаний и навыков в области организации процесса непрерывной интеграции и развертывания моделей интеллектуальных систем.

В рамках дисциплины у обучающихся формируются базовые представления и знания о жизненном цикле проектирования, реализации, валидации, развертывания и мониторинга моделей искусственного интеллекта, основаном на лучших практиках автоматизации процессов в соответствии с методологией MLOps.

На практических занятиях у обучающихся формируются навыки работы с инструментами контроля версий кода, данных и моделей, организации автоматизированных процессов управления экспериментами, интеграцией, валидацией, развертывания и мониторингом моделей искусственного интеллекта.

2. Планируемые результаты обучения по дисциплине (модулю).

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

ОПК-5 - Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем;

ОПК-6 - Способен разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования;

ПК-3 - Способен спроектировать, разработать, обучить, оценить и развернуть модели искусственного интеллекта в соответствии с методологией MLOps.

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

Уметь:

- использовать системы контроля версий для версионирования кода обучения моделей, данных и признаков, конфигурации виртуальных сред;

- использовать Apache Airflow для организации воспроизводимого конвейера машинного обучения;

- использовать инструменты Prometheus и Grafana для организации мониторинга работы моделей искусственного интеллекта.

- проектировать и реализовывать компоненты интеграции моделей искусственного интеллекта используя современные архитектуры;

- использовать Docker и K8s для развертывания компонентов интеллектуальных систем.

Знать:

- особенности постановки бизнес-целей разработки ML-решений;
- этапы жизненного цикла разработки ML-решений;
- основные концепции методологии MLOps;
- основы принципы работы систем контроля версий для артефактов моделирования искусственного интеллекта;
- принципы и подходы к управлению экспериментами;
- методы и стратегии валидации моделей искусственного интеллекта;
- особенности подходов к обработке;
- подходом и методов поиска отклонений и сдвигов в данных.

Владеть:

- работы с инструментами автоматизации настройки виртуальных сред развертывания и выполнения моделей искусственного интеллекта;
- навыком организации процессов автоматического сбора и подготовки данных для обучения в AirFlow;
- навыком организации процессов разработки моделей с использованием MLFlow;
- навык организации конвейнеров для автоматического анализа модели и подбора порогов;
- навык организации процессов развертывания и обновления моделей искусственного интеллекта.

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	Сем. №3

Контактная работа при проведении учебных занятий (всего):	36	36
В том числе:		
Занятия лекционного типа	18	18
Занятия семинарского типа	18	18

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 108 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	<p>Введение в воспроизводимые и масштабируемые процессы машинного обучения (ML).</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в организацию сред разработки для машинного обучения; - воспроизводимость и масштабируемость исследований; - виртуальные окружения, виды и особенности; - бизнес-метрики, цели и приоритеты; - типовой ML проект, типовые ошибки.
2	<p>Жизненный цикл разработки ML-решения.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в понятие MLOps, архитектура MLOps; - анализ требований, формирование целей; - сбор и подготовка данных; - машинное обучение и разработка моделей ИИ; - развертывание в различных виртуальных средах; - мониторинг; - обзор инструментов MLOps, облачные платформы.
3	<p>Хранение и контроль версий.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в проблематику хранения и версионирования кода и моделей ИИ;

№ п/п	Тематика лекционных занятий / краткое содержание
	<ul style="list-style-type: none"> - версионирование данных, обзор Data Version Control; - трекинг данных, восстановление данных, обеспечение воспроизводимости и масштабируемости с помощью конвейеров данных; - параметры, графики и метрики в DVC; - работа с FeatureStore; - обзор Github, Gitlab, Gitea для версионирования кода; - подход «Инфраструктура как код».
4	<p>Подходы к работе с данными на каждом этапе разработки ML-решений.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в работу с данными в виде таблиц, текста, картинок, аудио и видео, особенности разметки; - обзор особенностей работы с Pandas, SQL и NoSQL, Apache Spark, Hive в автоматизированных процессах; - обзор инструментов для планирования и выполнения задач по обработке данных; - организация процессов сбора и подготовки данных для обучения в AirFlow.
5	<p>Управление экспериментами.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в понятие эксперимента, подходы к управлению экспериментами; - технологии и инструменты для управления экспериментами; - управления экспериментами в MLFlow; - организация процессов разработки моделей с использованием MLFlow; - упаковка моделей; - организация репозитория и реестров моделей ИИ.
6	<p>Валидация моделей ИИ.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в стратегии валидации моделей; - обучающая, тестовая и валидационные выборки; - кросс-валидации, временная кросс-валидация; - анализ моделей ИИ, анализ метрик и выбор порогов, анализ производительности, анализ ошибок; - организация конвейеров для автоматического анализа модели и подбора порогов.
7	<p>Интеграция моделей ИИ.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в распространенные модели интеграции моделей ИИ; - архитектурные особенности современных систем, распределенные системы; - монолиты, модулиты, сервисы и микросервисы; - платформы, языки программирования и фреймворки; - особенности интеграции на облачных платформах; - модели пакетной обработки и обработки в реальном времени и их особенности интеграции; - связь жизненного цикла разработки моделей ИИ с жизненным циклом разработки ПО; - автоматизация интеграции моделей ИИ в рамках MLOps, обзор инструментов и практик.
8	<p>Развертывание моделей ИИ.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - особенности современных технологий развертывания, технологии, DevOps и MLOps практики; - обзор технологий контейнеризации, Docker, K8s; - особенности развертывания интеллектуальных систем пакетной обработки, AirFlow; - особенности развертывания интеллектуальных систем обработки в реальном времени, Kafka; - обновление моделей ИИ.
9	<p>Мониторинг.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - введение в сбор и визуализация метрик; - инструменты сбора и визуализации метрик Prometheus и Grafana; - анализ работы моделей, поиск отклонений и сдвигов в данных;

№ п/п	Тематика лекционных занятий / краткое содержание
	- оповещения, процессы обработки оповещений.

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Постановка задачи и бизнес-цели. В результате выполнения практической работы студент знакомится с постановкой целей разработки ML-решений, установкой приоритетов и формированием бизнес-целей.
2	Версионирование кода. В результате выполнения практической работы студент получает навык работы с инструментами версионирования кода – Gitlab и Gitea.
3	Версионирование данных. В результате выполнения практической работы студент получает навык версионирования данных и признаков с помощью инструментов Data Version Control и FeatureStore.
4	Инфраструктура как код. В результате выполнения практической работы студент получает навык работы с инструментами автоматизации настройки виртуальных сред развертывания и выполнения моделей ИИ.
5	Работа с данными. В результате выполнения практической работы студент получает навык организации процессов сбора и подготовки данных для обучения в AirFlow.
6	Управление экспериментами в MLFlow. В результате выполнения практической работы студент получает навык организации процессов разработки моделей с использованием MLFlow.
7	Управление валидацией. В результате выполнения практической работы студент получает навык организации конвейнеров для автоматического анализа модели и подбора порогов.
8	Управление интеграцией. В результате выполнения практической работы студент получает навык организации процессов интеграции моделей ИИ.
9	Управление развертыванием. В результате выполнения практической работы студент получает навык организации процессов развертывания и обновления моделей ИИ.
10	Мониторинг. В результате выполнения практической работы студент получает навык сбора и визуализации метрик с помощью Prometheus и Grafana.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение рекомендованной литературы.
2	Подготовка к практическим занятиям.
3	Выполнение курсовой работы.
4	Подготовка к промежуточной аттестации.

4.4. Примерный перечень тем курсовых работ

1. Организация процессов версионирования кода обучения моделей ИИ на платформе Gitlab.
2. Организация процессов версионирования инфраструктурного кода и конфигураций на платформе Gitea.
3. Организация процессов версионирования данных с использованием Data Version Control.
4. Организация процессов версионирования признаков с использованием FeatureStore.
5. Организация процессов сбора и подготовки данных для пакетной обработки с использованием AirFlow.
6. Организация процессов сбора и подготовки данных для обработки данных в реальном времени с использованием Kafka и Spark Stream.
7. Обзор свободно-распространяемых MLOps решений.
8. Особенности автоматизации процессов для пакетной обработки в рамках управления экспериментами в MLFlow.
9. Особенности автоматизации процессов для обработки данных в реальном времени в рамках управления экспериментами в MLFlow.
10. Docker для процессов моделирования ИИ.
11. Визуализация и анализ метрик моделей ИИ.

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

№ п/п	Библиографическое описание	Место доступа
1	Хапке, Х. Разработка конвейеров машинного обучения : руководство / Х. Хапке, К. Нельсон ; перевод с английского Н. Б. Желновой. — Москва : ДМК Пресс, 2021. — 346 с. — ISBN 978-5-97060-886-9	https://e.lanbook.com/book/241088 (дата обращения: 1.11.2022). – Текст : электронный
2	Баймуратов, И. Р. Методы автоматизации машинного обучения : учебное пособие / И. Р. Баймуратов. — Санкт-Петербург : НИУ ИТМО, 2020. — 40 с.	https://e.lanbook.com/book/190871 (дата обращения: 1.11.2022). – Текст : электронный
3	Харенслак, Б. Apache Airflow и конвейеры обработки данных / Б. Харенслак, Р. Д. де ; перевод с английского Д. А. Беликова. — Москва : ДМК	https://e.lanbook.com/book/241133 (дата обращения: 1.11.2022). – Текст : электронный

	Пресс, 2022. — 502 с. — ISBN 978-5-97060-970-5	
4	Сейерс, Э. Х. Docker на практике / Э. Х. Сейерс, А. Милл ; перевод с английского Д. А. Беликов. — Москва : ДМК Пресс, 2020. — 516 с. — ISBN 978-5-97060-772-5	https://e.lanbook.com/book/131719 (дата обращения: 1.11.2022). — Текст : электронный
5	Кочер, П. С. Микросервисы и контейнеры Docker : руководство / П. С. Кочер ; перевод с английского А. Н. Киселева. — Москва : ДМК Пресс, 2019. — 240 с. — ISBN 978-5-97060-739-8	https://e.lanbook.com/book/123710 (дата обращения: 1.11.2022). — Текст : электронный
6	Пселтис, Э. Д. Поточковая обработка данных. Конвейер реального времени / Э. Д. Пселтис ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2018. — 218 с. — ISBN 978-5-97060-606-3	https://e.lanbook.com/book/105840 (дата обращения: 1.11.2022). — Текст : электронный

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

Электронно-библиотечная система Научно-технической библиотеки РУТ(МИИТ) (<http://library.miit.ru/>)

Электронно-библиотечная система издательства «Лань» (<http://e.lanbook.com/>)

Электронно-библиотечная система ibooks.ru (<http://ibooks.ru/>)

Открытые лекции (<https://sphere.vk.company/materials/video/#19>)

Открытые лекции (<https://ods.ai/tracks/ml-in-production-spring-22>)

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

Браузер Microsoft Internet Explorer или его аналоги

Пакет офисных программ Microsoft Office или его аналоги

Python 3.6 и выше

Anaconda

Data Version Control

FeatureStore

Apache Airflow

MLflow

GitLab

Prometheus

Grafana

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

Учебные аудитории для проведения учебных занятий, оснащенные компьютерной техникой и наборами демонстрационного оборудования.

Для практических занятий – наличие персональных компьютеров вычислительного класса.

9. Форма промежуточной аттестации:

Зачет в 3 семестре.

Курсовая работа в 3 семестре.

10. Оценочные материалы.

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

Авторы:

заведующий кафедрой, доцент, к.н.
кафедры «Цифровые технологии
управления транспортными
процессами»

В.Е. Нутович

старший преподаватель кафедры
«Цифровые технологии управления
транспортными процессами»

Е.А. Заманов

Согласовано:

Заведующий кафедрой ЦТУТП

В.Е. Нутович

Председатель учебно-методической
комиссии

Н.А. Андриянова