МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ)

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Продвинутая ML-аналитика (оптимизация и тюнинг)

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): Цифровая инженерия транспортных процессов

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 ФГАОУ ВО РУТ (МИИТ)
ID подписи: 937226
Подписал: руководитель образовательной программы Проневич Ольга Борисовна
Дата: 08.09.2026

1. Общие сведения о дисциплине (модуле).

Целями освоения учебной дисциплины являются формирование у обучающихся системных знаний и практических навыков в области продвинутой машинно-обучаемой аналитики, оптимизации и тонкой настройки моделей машинного обучения для решения прикладных задач транспортной отрасли. В рамках курса рассматриваются алгоритмы обучения с учителем и без учителя, методы оптимизации гиперпараметров моделей, подходы к интерпретации результатов работы моделей, а также методы автоматического машинного обучения (AutoML) и их адаптация к специфике транспортных задач. Особое внимание уделяется планированию экспериментов, формулировке статистических гипотез, выбору метрик для оценки эффекта воздействия, расчету размера выборки и интерпретации результатов экспериментов применительно к задачам прогнозирования, кластеризации и обнаружения аномалий в транспортных данных.

Задачами освоения дисциплины являются:

Формирование теоретических знаний об алгоритмах обучения с учителем и без учителя, методах оптимизации гиперпараметров и автоматического машинного обучения.

Освоение методов анализа и содержательной интерпретации результатов работы моделей машинного обучения.

Развитие практических навыков тонкой настройки моделей и адаптации алгоритмов к специфике транспортных задач.

Формирование компетенций по планированию экспериментов, формулировке статистических гипотез и выбору метрик для оценки эффекта воздействия.

Освоение методов расчета размера выборки для достижения заданной мощности теста.

Приобретение навыков обоснования способов и вариантов применения алгоритмов автоматического машинного обучения в задачах искусственного интеллекта.

Дисциплина предназначена для формирования компетенций, необходимых для эффективного применения и оптимизации моделей машинного обучения в транспортных организациях.

2. Планируемые результаты обучения по дисциплине (модулю).

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

МП-7 - Способен использовать инструменты ИИ для быстрого прототипирования, валидации идей и тестирования спроса;

ОПК-3 - Способен использовать современные информационные технологии и программно-аппаратные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.

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

Знать:

Алгоритмы обучения с учителем для регрессии и классификации (линейные методы, ансамблевые методы, методы на опорных векторах) и алгоритмы обучения без учителя (методы кластеризации и поиска аномалий).

Методы оптимизации гиперпараметров моделей (Grid Search, Random Search, Bayesian Optimization, Optuna) и подходы к интерпретации результатов работы моделей (SHAP, LIME, важность признаков).

Основные принципы планирования экспериментов: рандомизация, контрольные группы, мощность теста, а также методы расчета размера выборки для достижения заданной мощности.

Принципы автоматического машинного обучения (AutoML), включая автоматический выбор модели и настройку гиперпараметров, и подходы к их адаптации для специфики транспортных задач.

Уметь:

Анализировать и содержательно интерпретировать результаты работы моделей машинного обучения, включая оценку важности признаков и объяснение отдельных предсказаний.

Обосновывать способы и варианты применения алгоритмов автоматического машинного обучения в задачах искусственного интеллекта, включая их преобразование и адаптацию к специфике транспортных задач.

Формулировать статистические гипотезы (H?, H?), выбирать метрики для оценки эффекта воздействия и планировать эксперименты с учетом принципов рандомизации и контрольных групп.

Применять методы оптимизации гиперпараметров для тонкой настройки моделей машинного обучения с использованием современных фреймворков (Optuna, Hyperopt).

Владеть:

Навыками оптимизации гиперпараметров моделей с использованием различных методов (Grid Search, Random Search, Bayesian Optimization) и интерпретации результатов моделирования.

Навыками применения методов интерпретации моделей (SHAP, LIME) для объяснения предсказаний и выявления ключевых факторов, влияющих на результат.

Техниками планирования и проведения экспериментов, расчета размера выборки, оценки мощности теста и документирования результатов экспериментальных исследований.

Навыками работы с инструментами AutoML для автоматического выбора моделей и настройки гиперпараметров применительно к транспортным задачам.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 2 з.е. (72 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 №3
Контактная работа при проведении учебных занятий (всего):	32	32
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского типа	16	16

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 40 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. Введение в продвинутую ML-аналитику. Рассматриваемые вопросы: - Понятие продвинутой ML-аналитики. - Роль оптимизации и тюнинга в жизненном цикле моделей. - Особенности транспортных задач для машинного обучения. - Основные вызовы и тренды. - Обзор инструментов для продвинутой аналитики.
2	Тема 2. Алгоритмы обучения с учителем и без учителя. Рассматриваемые вопросы: - Линейные методы: линейная и логистическая регрессия. - Методы на опорных векторах (SVM). - Ансамблевые методы: Random Forest, Gradient Boosting. - Методы кластеризации: K-means, DBSCAN, иерархическая кластеризация. - Методы поиска аномалий: Isolation Forest, Local Outlier Factor. - Выбор алгоритма для транспортных задач.
3	Тема 3. Методы оптимизации гиперпараметров. Рассматриваемые вопросы: - Понятие гиперпараметров и их влияние на модель. - Grid Search и Random Search. - Bayesian Optimization. - Optuna и другие современные фреймворки. - Выбор метода оптимизации для транспортных задач.
4	Тема 4. Интерпретация результатов работы моделей. Рассматриваемые вопросы: - Важность признаков. - SHAP и LIME для интерпретации предсказаний. - Анализ ошибок модели. - Визуализация результатов. - Формирование выводов и рекомендаций.
5	Тема 5. Автоматическое машинное обучение (AutoML). Рассматриваемые вопросы: - Понятие и задачи AutoML. - Обзор инструментов AutoML. - Автоматический выбор модели. - Автоматическая настройка гиперпараметров. - Адаптация AutoML к специфике транспортных задач.
6	Тема 6. Планирование экспериментов в машинном обучении. Рассматриваемые вопросы: - Основные принципы планирования экспериментов. - Рандомизация и контрольные группы. - Формулировка статистических гипотез (H?, H?). - Выбор метрик для оценки эффекта воздействия. - Методы расчета размера выборки.
7	Тема 7. Оценка мощности теста и интерпретация экспериментов. Рассматриваемые вопросы:

№ п/п	Тематика лекционных занятий / краткое содержание
	- Понятие мощности теста. - Факторы, влияющие на мощность. - Методы расчета размера выборки для заданной мощности. - Интерпретация результатов экспериментов. - Документирование экспериментальных исследований.
8	Тема 8. Статистические гипотезы и метрики в ML-экспериментах. Рассматриваемые вопросы: - Типы статистических гипотез и их применение в ML. - Выбор метрик для оценки эффекта воздействия. - Связь метрик с бизнес-показателями. - Ошибки первого и второго рода. - Практические аспекты проверки гипотез в транспортных задачах.

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Тема 1. Введение в Python для продвинутой ML-аналитики. Рассматриваемые вопросы: - Обзор библиотек для ML. - Настройка среды разработки. - Загрузка и подготовка транспортных данных. - Базовые операции с данными. - Первичный анализ данных.
2	Тема 2. Реализация линейных моделей для транспортных задач. Рассматриваемые вопросы: - Линейная регрессия для прогнозирования. - Логистическая регрессия для классификации. - Оценка качества моделей. - Интерпретация коэффициентов. - Применение к транспортным данным.
3	Тема 3. Реализация методов на опорных векторах (SVM). Рассматриваемые вопросы: - Линейный SVM и SVM с ядровыми функциями. - Выбор параметров SVM. - Применение к задачам классификации на транспорте. - Оценка качества.
4	Тема 4. Реализация ансамблевых методов. Рассматриваемые вопросы: - Random Forest для регрессии и классификации. - Gradient Boosting (XGBoost, LightGBM). - Настройка базовых параметров. - Сравнение ансамблевых методов. - Применение к транспортным задачам.
5	Тема 5. Методы кластеризации транспортных данных. Рассматриваемые вопросы: - K-means кластеризация. - DBSCAN для данных со сложной структура. Иерархическая кластеризация.

№ п/п	Тематика практических занятий/краткое содержание
	- Оценка качества кластеризации. - Интерпретация кластеров.
6	Тема 6. Обнаружение аномалий в транспортных данных. Рассматриваемые вопросы: - Isolation Forest. - Local Outlier Factor. - Применение к данным GPS/ГЛОНАСС. - Обнаружение выбросов в логистических данных. - Визуализация аномалий.
7	Тема 7. Grid Search для оптимизации гиперпараметров. Рассматриваемые вопросы: - Определение пространства гиперпараметров. - Настройка Grid Search. - Кросс-валидация в Grid Search. - Оценка результатов. - Применение к транспортным моделям.
8	Тема 8. Random Search для оптимизации гиперпараметров. Рассматриваемые вопросы: - Настройка Random Search. - Сравнение Grid Search и Random Search. - Выбор оптимальных параметров. - Применение к транспортным задачам. - Анализ результатов.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение литературы и документации по алгоритмам машинного обучения (линейные методы, ансамблевые методы, SVM).
2	Изучение методов оптимизации гиперпараметров и фреймворков.
3	Самостоятельное изучение методов интерпретации моделей (SHAP, LIME).
4	Подготовка к промежуточной аттестации.
5	Подготовка к промежуточной аттестации.
6	Подготовка к текущему контролю.

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

№ п/п	Библиографическое описание	Место доступа
1	Пимахина, Е. В. Современные подходы к изучению сепсиса в практике анестезиолога-реаниматолога : учебное пособие / Е. В.	https://e.lanbook.com/book/523988

	Пимахина. — Рязань : РязГМУ, 2026. — 87 с. — ISBN 978-5-8423-0308-3.	
2	Methods in medical genetics: tutorial : учебное пособие / R. N. Mustafin, G. F. Gimalova, I. R. Gilyazova [и др.]. — Уфа : БГМУ, 2025. — 102 с.	https://e.lanbook.com/book/521965

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

http://library.miit.ru/ - электронно-библиотечная система научно-технической библиотеки МИИТ

https://superset.apache.org/user-docs/intro/#documentation – документация Apache Superset

https://www.postgresql.org/ - официальная страница PostgreSQL

https://yandex.cloud/ru/docs/datalens/ - официальная документация Yandex Datalens

https://docs.docker.com/ - официальная документация Docker

https://www.metabase.com/docs/latest/ - официальная документация Metabase

https://tableau.pro/ - официальный сайт Tableau с лучшими практиками проектирования дашбордов

https://revealthedata.com/blog/ - Блог о визуализации данных и BI-системах

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

MS Office (Excel, Word, PowerPoint)
Apache Superset
Yandex Datalens
Metabase
PostgreSQL
Docker
Apache Airflow
Браузер Chrome / Яндекс браузер
Draw.io

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

Компьютер преподавателя с доступом к сети Интернет

Компьютеры студентов с установленным Python-окружением и табличным процессором

Доступ к учебной базе событийных данных и экспериментальным выборкам

Доступ к PostgreSQL и BI-системе для визуализации результатов

Доступ к облачной виртуальной машине или Docker-контейнерам для воспроизводимого анализа

Проектор и экран для проектора

Маркерная доска

9. Форма промежуточной аттестации:

Зачет в 3 семестре.

10. Оценочные материалы.

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

Авторы:

руководитель образовательной программы О.Б. Проневич

Согласовано:

Директор Д.В. Паринов

Руководитель образовательной программы О.Б. Проневич

Председатель учебно-методической комиссии Д.В. Паринов