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)

Прикладные методы анализа данных на транспорте

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).

Цели освоения учебной дисциплины:

Формирование у обучающихся профессиональных компетенций в области анализа данных на транспорте, включающих способность извлекать, обрабатывать и интерпретировать большие массивы пространственно-временной информации для решения прикладных задач управления мобильностью, логистики и обеспечения безопасности дорожного движения.

Задачи освоения дисциплины:

Освоение принципов проектирования иерархии данных и навигации в панелях мониторинга для эффективного представления результатов анализа.

Изучение принципов юзабилити и визуального сторителлинга для создания информативных и удобных дашбордов, адаптированных под различные группы пользователей.

Освоение методов анализа данных на транспорте, включая дескриптивную аналитику транспортных потоков, геопространственный анализ, анализ временных рядов и прогнозирование.

Формирование практических навыков применения методов машинного обучения (классификация, кластеризация, регрессия) для решения прикладных транспортных задач.

Освоение методов оптимизации маршрутизации и логистических процессов с использованием алгоритмов поиска кратчайших путей и решения задач VRP.

Приобретение навыков имитационного моделирования транспортных процессов на микроуровне для оценки эффективности управленческих решений.

Развитие компетенций по интерпретации результатов анализа и моделей машинного обучения (SHAP) для обоснования транспортных решений перед заинтересованными сторонами.

2. Планируемые результаты обучения по дисциплине (модулю).

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

ППК-ДАЗ - Способен проводить аналитические исследования с применением технологий больших данных.

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

Знать:

Специфику транспортных данных: пространственно-временную природу, разреженность, проблему пропущенных значений в датчиках, источники данных (детекторы транспорта, GPS-треки, данные валидации проездных, видеопотоки, расписания в формате GTFS).

Фундаментальную диаграмму транспортного потока (скорость, плотность, интенсивность), статистические метрики загруженности дорог (LOS, коэффициент заторов) и методы очистки «сырых» данных с детекторов транспорта.

Принципы построения графов дорожных сетей (OSMnx, NetworkX): понятие ребра, узла, веса (время/расстояние), методы геокодирования и преобразования координат, построение матриц корреспонденций (OD-matrix).

Уметь:

Настраивать среду разработки для транспортного аналитика (Jupyter, Pandas, GeoPandas), выполнять импорт открытых транспортных данных (расписания в формате GTFS, данные велопроката, статистика ДТП) и оценивать качество данных.

Выполнять анализ аварийности и безопасности движения: загружать статистику ДТП, рассчитывать плотность очагов аварийности методом ядерной оценки плотности (KDE), определять «чёрные точки» на карте.

Визуализировать интенсивность транспортных потоков: строить тепловые карты и картограммы загрузки дорог, создавать анимированные графики движения по GPS-трекам, визуализировать матрицы корреспонденций потоковыми линиями.

Владеть:

Навыками работы с Python-стеком для транспортного аналитика: Pandas, GeoPandas, Matplotlib, Seaborn, Plotly, Folium, OSMnx, NetworkX.

Навыками построения прогнозных моделей временных рядов (SARIMA, Holt-Winters) и оценки их точности для прогнозирования пассажиропотока и интенсивности движения.

Навыками применения методов машинного обучения для решения транспортных задач: классификация (Random Forest, XGBoost), кластеризация (K-Means, DBSCAN), регрессия (линейная, градиентный бустинг).

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 №2
Контактная работа при проведении учебных занятий (всего):	48	48
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского 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 24 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 анализ данных на транспорте: специфика больших данных Рассматриваются вопросы: Источники данных на транспорте: детекторы транспорта, GPS-треки, валидации электронных проездных, данные видеопотоков и расписаний. Понятие «Intelligent Transportation Systems» (ITS). Роль Data Science в управлении мобильностью. Особенности транспортных данных: пространственно-временная природа, разреженность, телематика, проблема пропущенных значений в датчиках.
2	Тема 2. Дескриптивная аналитика и базовое моделирование транспортных потоков Рассматриваются вопросы:

№ п/п	Тематика лекционных занятий / краткое содержание
	Фундаментальная диаграмма транспортного потока (скорость, плотность, интенсивность). Статистические метрики для оценки загруженности дорог: уровень обслуживания (LOS), коэффициент заторов. Методы очистки и фильтрации «сырых» данных с детекторов транспорта (алгоритмы отсеивания шума).
3	Тема 3. Геопространственный анализ транспортных сетей Рассматриваются вопросы: Модели графов дорожных сетей (OSMnx, NetworkX). Понятие ребра, узла, веса (время/расстояние). Геокодирование и преобразование координат. Построение матриц корреспонденций (OD-matrix): откуда и куда едут пассажиры или грузы.
4	Тема 4. Анализ временных рядов в транспортных системах Рассматриваются вопросы: Декомпозиция транспортного спроса: тренд, сезонность (утренний/вечерний час пик, день недели), остатки. Прогнозирование интенсивности движения методами SARIMA и экспоненциального сглаживания (Holt-Winters). Детекция аномалий в транспортных потоках (ДТП, внезапные пробки) с помощью статистических критериев.
5	Тема 5. Прикладные задачи классификации и кластеризации на транспорте Рассматриваются вопросы: Классификация типов транспортных средств по данным с индуктивных петель и акселерометров смартфонов. Сегментация пассажиров транспортной системы (RFM-анализ по истории поездок). Кластеризация остановочных пунктов по пассажиропотоку для оптимизации маршрутной сети (K-Means, DBSCAN).
6	Тема 6. Регрессионный анализ и предиктивное моделирование в логистике Рассматриваются вопросы: Факторы, влияющие на время доставки «последней мили» (погода, загруженность курьера, расстояние). Построение линейной и нелинейной (Random Forest, Gradient Boosting) регрессии. Прогнозирование износа узлов подвижного состава (предиктивное техническое обслуживание) по телеметрическим данным.
7	Тема 7. Маршрутизация и оптимизация логистических процессов Рассматриваются вопросы: Классическая задача коммивояжера (TSP) и задача маршрутизации транспорта (VRP) с ограничениями по грузоподъемности и времени. Алгоритм Дейкстры и A* для навигации. Поиск кратчайшего пути с учетом прогноза пробок. Моделирование пассажирских пересадок и построение оптимальных расписаний.
8	Тема 8. Имитационное моделирование и визуальная коммуникация результатов Рассматриваются вопросы: Агентное моделирование пешеходных потоков и перекрестков (микро-уровень). Основы построения транспортных дашбордов (Dash, Streamlit): визуализация интенсивности на карте. Интерпретация моделей машинного обучения (SHAP) для обоснования транспортных решений

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Практика 1. Python-стек для транспортного аналитика: первичная загрузка данных Рассматриваются вопросы: Настройка среды (Jupyter, Pandas, GeoPandas). Импорт открытых данных (расписания в формате GTFS, данные велопроката). Оценка качества данных: поиск дубликатов в журнале транзакций оплаты проезда.
2	Практика 2. Анализ аварийности и безопасности движения Рассматриваются вопросы: Загрузка статистики ДТП (формат CSV с координатами). Расчет плотности очагов аварийности с помощью ядерной оценки плотности (KDE). Определение «черных точек» (Hotspot Analysis) на карте города.
3	Практика 3. Визуализация интенсивности транспортных потоков Рассматриваются вопросы: Построение тепловых карт и картограмм загрузки дорог почасово. Создание анимированного графика движения автобусов по GPS-трекам во времени. Визуализация матрицы корреспонденций с помощью потоковых линий.
4	Практика 4. Анализ транзакций бесконтактной оплаты проезда Рассматриваются вопросы: Агрегация данных валидации (количество входов в метро/автобус) по часам. Профилирование пассажира: определение домашней станции по частоте и времени первой/последней поездки. Оценка доли пересадок и времени ожидания между рейсами.
5	Практика 5. Построение и анализ графа улично-дорожной сети Рассматриваются вопросы: Извлечение графа дорог района города через библиотеку OSMnx. Расчет центральности узлов (Betweenness Centrality). Выявление критически важных перекрестков, разрушение которых парализует движение.
6	Практика 6. Временные ряды: прогнозирование пассажиропотока Рассматриваются вопросы: Подготовка данных: ресемплинг количества пассажиров по дням. Построение модели SARIMA для прогноза спроса на перевозки на следующую неделю. Валидация модели и оценка точности MAPE.
7	Практика 7. Детекция инцидентов на основе скорости транспортного потока Рассматриваются вопросы: Работа с данными дорожных датчиков (скорость, занимаемость полосы). Применение экспоненциального сглаживания для выявления резких скачков. Реализация алгоритма Калифорнийского теста для автоматического обнаружения ДТП.
8	Практика 8. Классификация режимов передвижения (Transport Mode Detection) Рассматриваются вопросы: Работа с данными акселерометра и GPS со смартфона. Feature Engineering: выделение признаков (средняя скорость, ускорение, извилистость траектории). Построение классификатора (Random Forest) для различения: пешком, велосипед, авто, электричка.
9	Практика 9. Сегментация пользователей каршеринга Рассматриваются вопросы: Создание RFM-матрицы: Recency (давность), Frequency (частота), Mileage (пробег). Стандартизация признаков и кластеризация K-Means. Психологический портрет кластеров: «туристы», «маятниковые мигранты», «драйверы выходного дня».
10	Практика 10. Предиктивное обслуживание узлов грузового вагона Рассматриваются вопросы: Анализ логов телеметрии с буксовых датчиков (температура, вибрация).

№ п/п	Тематика практических занятий/краткое содержание
	Очистка выбросов и интерполяция пропусков. Бинарная классификация: прогноз отказа буксового узла в ближайшие 30 часов на основе градиентного бустинга (XGBoost).
11	Практика 11. Прогнозирование времени прибытия такси (ETA) Рассматриваются вопросы: Формирование трейнового датасета из истории поездок такси. Признаки: дистанция по прямой, дистанция по дороге, час пик, день недели, погода. Построение регрессионной модели LightGBM и анализ важности признаков.
12	Практика 12. Решение задачи оптимизации маршрутов (VRP) Рассматриваются вопросы: Постановка задачи: развоз товаров со склада по N точкам ограниченным числом курьеров. Использование библиотеки OR-Tools для поиска оптимального плана поездок. Сравнение оптимального маршрута с фактическим (по трекам водителей).
13	Практика 13. Моделирование остановочного пункта (Агентное моделирование) Рассматриваются вопросы: Построение простейшей имитационной модели посадки/высадки пассажиров в автобус. Узкие места: анализ времени блокировки дверей и задержки отправления при разных пассажиропотоках.
14	Практика 14. Анализ устойчивости расписания движения поездов Рассматриваются вопросы: Расчет пунктуальности движения (процент опозданий > 5 мин). Построение графика «время-путь» (Time-Space diagram) для ниток графика. Визуальное выявление конфликтов скрещения и обгона на однопутной линии.
15	Практика 15. A/B тестирование транспортных реформ Рассматриваются вопросы: Оценка эффективности выделенной полосы для автобусов. Сравнение среднего времени проезда «до» и «после» с использованием t-теста. Проверка гипотез: действительно ли изменилась скорость сообщения на смежных полосах.
16	Практика 16. Выявление скрытых закономерностей в ДТП (Apriori-анализ) Рассматриваются вопросы: Дискретизация данных о ДТП: время суток, погодные условия, тип столкновения, освещение. Применение алгоритма Apriori для поиска ассоциативных правил (например, «Гололёд + Неосвещённый переход -> Наезд на пешехода»). Оценка значимости правил через метрики Support, Confidence, Lift.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение литературы по BI-аналитике.
2	Анализ существующих дашбордов на примерах.
3	Сбор и подготовка данных для своего проекта
4	Разработка собственного дашборда по транспортной тематике.
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	Оськин, Д. А. Интеллектуальный анализ данных : учебник для вузов / Д. А. Оськин, Н. Г. Левченко, Н. А. Седова. — Санкт-Петербург : Лань, 2026. — 228 с. — ISBN 978-5-507-56021-9.	https://e.lanbook.com/book/522509
2	Вольфсон, М. Б. Анализ данных : учебно-методическое пособие / М. Б. Вольфсон. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2023. — 69 с.	https://e.lanbook.com/book/381533

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 во 2 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