МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ)

АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Инженерия данных на транспорте

Направление подготовки:	09.03.01 – Информатика и вычислительная техника
Направленность (профиль):	Цифровая инженерия транспортных процессов
Форма обучения:	Очная

Общие сведения о дисциплине (модуле).

Целями освоения учебной дисциплины являются формирование у обучающихся системных знаний и практических навыков в области инженерии данных для решения задач транспортной отрасли. В рамках курса рассматриваются жизненный цикл данных в транспортных системах, методы сбора, хранения и предобработки структурированных и неструктурированных данных, а также основы построения конвейеров данных (ETL/ELT) для последующего применения моделей машинного обучения и эконометрического анализа. Особое внимание уделяется работе с транспортными логистическими потоками, телематическими данными, базами данных ГЛОНАСС/GPS, а также автоматизации процессов подготовки признаков для прогнозных моделей.

Задачами освоения дисциплины являются:

- Формирование теоретических знаний об источниках данных на транспорте (GPS-треки, логистические системы, АСУ, метеоданные) и их специфике.

- Освоение методов первичного анализа данных (EDA) для выявления структуры, пропусков, выбросов и аномалий в транспортных данных.

- Развитие практических навыков преобразования и нормализации данных для обеспечения их совместимости с различными моделями машинного обучения.

- Формирование компетенций по автоматизации процессов очистки и подготовки признаков с использованием инструментов Python.

- Освоение критериев выбора модели в зависимости от типа решаемой задачи (классификация, регрессия, кластеризация) и характера транспортных данных.

- Приобретение навыков оценки качества моделей и диагностики переобучения/недообучения применительно к задачам прогнозирования на транспорте.

Дисциплина предназначена для формирования компетенций, необходимых для эффективной работы с данными на этапе их подготовки и первичного моделирования в транспортных организациях.

Общая трудоемкость дисциплины (модуля) составляет 2 з.е. (72 академических часа(ов).