

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
38.03.01 Экономика,
утвержденной директором РУТ (МИИТ)
Игольниковым Б.В.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Основы анализа данных

Направление подготовки: 38.03.01 Экономика

Направленность (профиль): Экономика и инженерия транспортных систем. Программа двойного диплома с Высшей школой экономики

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:
Подписал:
Дата: 28.02.2025

1. Общие сведения о дисциплине (модуле).

Курс от простого к сложному погружает слушателей в теорию и практические аспекты работы с большими массивами информации. В рамках курса на простом и понятном языке будет рассказано как в современных условиях производится обобщение статистических данных и поиск скрытых закономерностей, а также как эти закономерности используются при принятии управленческих решений и повышении эффективности работы предприятия. Основная цель данной дисциплины – формирование у студентов основных теоретических знаний и практических навыков в области анализа данных, применении наиболее популярных методов статистического анализа и прогностической аналитики.

2. Планируемые результаты обучения по дисциплине (модулю).

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

ОПК-2 - Способен осуществлять сбор, обработку и статистический анализ данных, необходимых для решения поставленных экономических задач;

УК-1 - Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач.

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

Владеть:

языками программирования, предназначенными для анализа и визуализации данных,

инструментами инженерии данных,

статистическими инструментами проверки гипотез о данных.

Уметь:

Выполнять профилирование данных и правильно определить какие методы анализа должны быть использованы в конкретной ситуации,

выбирать инструменты анализа данных в зависимости от их типов, объем,

собирать данные из открытых источников.

Знать:

Стратегии управления данными,

программное обеспечения, применяемые для анализа данных,

языки программирования, специализированные на анализе данных
Знание ландшафта альтернативных решений и особенностей их применения.

терминология в области анализа данных: программное обеспечение в области анализа данных, методы анализа.

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 №6
Контактная работа при проведении учебных занятий (всего):	34	34
В том числе:		
Занятия семинарского типа	34	34

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 38 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.

Не предусмотрено учебным планом

4.2. Занятия семинарского типа.

Лабораторные работы

№ п/п	Наименование лабораторных работ / краткое содержание
1	Введение в предметную область аналитики больших данных Рассматриваемые вопросы: <ul style="list-style-type: none">- данные, информация, знания – в чем отличия;- что есть большие данные. Что есть аналитика;- ожидания рынка vs результаты.
2	Введение в статистический анализ: разбираемся в основах Рассматриваемые вопросы: <ul style="list-style-type: none">- Основные понятия: генеральная совокупность и выборка.- Какие основные статистические показатели используются в аналитике и почему (рассмотрим на известных исторических примерах).- Что такое сводная таблица и как она используется в работе (рассмотрим на примере задачи финансового планирования).
3	Аналитика: сравнение групп данных Рассматриваемые вопросы: <ul style="list-style-type: none">- Для чего и где применяется.- Понятие значимости различий при сравнении групп (выборки) данных.- Зависимые и независимые выборки.- Таблицы перекрестной классификации.- Сравнение независимых выборок.- Сравнение зависимых выборок.- Анализ последовательностей.
4	Аналитика: анализ взаимосвязей Рассматриваемые вопросы: <ul style="list-style-type: none">- Зачем нужно анализировать связи в данных (примеры задач).- Зависимые и независимые переменные.- Статистические взаимосвязи и их значимость для анализа.- Корреляции.- Регрессия и её виды.- Факторный анализ.- Анализ надежности согласованности.- Понятие проклятия размерности.
5	Аналитика: классификация и что это даёт Рассматриваемые вопросы: <ul style="list-style-type: none">- Зачем нужна классификация (примеры задач).- Виды классификации.- Обзор используемых методов: деревья решений, случайный лес, наивный байесовский классификатор, линейная и логистическая регрессия.- Кластеризация (и чем принципиально отличается от классификации).- О чем важно помнить при выполнении кластеризации.
6	Аналитика: прогнозирование, хотим управлять будущим Рассматриваемые вопросы: <ul style="list-style-type: none">- Что есть прогнозирование, где и как используется.- Временные ряды.- Что такое сезонность и как её найти в данных.- Методы прогнозирования и их сравнение: линейная регрессия, экспоненциальное сглаживание, фильтр Кальмана, модель сквозящего среднего, авторегрессионная модель, ARIMA.

№ п/п	Наименование лабораторных работ / краткое содержание
7	<p>Машинное обучение: где и как применяется</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Зачем нужно машинное обучение, что скрывается под этим названием. - Методы машинного обучения для изучения и подготовки данных. - Примеры задач и методов их решения. - Обучение с подкреплением. - А/В тестирование.
8	<p>Построение визуализации и подготовка данных</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> - Решение практических задач по подготовке данных - Анализ качества - Расчетные показатели - Визуализация данных.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Работа с интернет-источниками и литературой
2	Настройка программного обеспечения
3	Изучение технической документации на программное обеспечение в области анализа данных и отслеживание обновления библиотек
4	Подготовка к промежуточной аттестации.

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

№ п/п	Библиографическое описание	Место доступа
1	<p>Основы интеллектуального анализа данных. Лабораторный практикум: учебное пособие Нестеров С.А. Издательство "Лань" , 2020</p>	https://reader.lanbook.com/book/130181#1
2	<p>Груздев, А. В. Прогнозное моделирование в IBM SPSS Statistics, R и Python: метод деревьев решений и случайный лес : руководство / А. В. Груздев. — Москва : ДМК Пресс, 2018. — 642 с. — ISBN 978-5-97060-539-4</p>	https://e.lanbook.com/book/123700
3	<p>Нестеров, С. А. Основы интеллектуального анализа данных. Лабораторный практикум : учебное пособие / С. А. Нестеров. — Санкт-</p>	https://e.lanbook.com/book/130181

	Петербург : Лань, 2020. — 40 с. — ISBN 978-5-8114-4509-7	
1	Прогнозное моделирование в IBM SPSS Statistics, R и Python: метод деревьев решений и случайный лес Груздев А.В. Издательство "ДМК Пресс" , 2018	https://reader.lanbook.com/book/123700#484

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).

Офисный пакет приложений – Microsoft Office

ПО для анализа данных Polymatica

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:

Зачет в 6 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

Н.В. Атамасов

Согласовано:

Председатель учебно-методической
комиссии

Д.В. Паринов