

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
01.03.02 Прикладная математика и информатика,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Анализ данных и временные ряды

Направление подготовки: 01.03.02 Прикладная математика и информатика

Направленность (профиль): Математическое моделирование и системный анализ

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 дисциплины (модуля) является:

- ознакомление студентов с элементами математической теории анализа данных и временных рядов.

Задачами освоения дисциплины (модуля) являются:

- формирование личности студента, развитие его интеллекта и умения логически и алгоритмически мыслить, формирование умений и навыков, необходимых при практическом применении теории комплексного анализа;

- обучение использованию в своей деятельности современных статистико-математические методов и моделей.

2. Планируемые результаты обучения по дисциплине (модулю).

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

ПК-4 - Уметь ставить цели создания системы, разрабатывать концепцию системы и требования к ней, выполнять декомпозицию требований к системе.

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

Знать:

- имеющиеся методы и средства решения прикладных задач;
- понятия и факты из области математических, а также других естественно-научных дисциплин.

Уметь:

- формулировать постановку задачи и излагать ее;
- применять на практике изученные методы и подходы.

Владеть:

- навыками классификации и выделения существенных подзадач при анализе данных;
- навыками сбора и обработки данных.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 3 з.е. (108 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 №7
Контактная работа при проведении учебных занятий (всего):	64	64
В том числе:		
Занятия лекционного типа	32	32
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 44 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	Доверительные интервалы Рассматриваемые вопросы: - квантили и интерквартильный размах; - предсказательный интервал; - доверительный интервал; - минимально необходимые объемы выборок; - минимальный обнаружимый эффект.
2	Проверка гипотез Рассматриваемые вопросы: - достигаемый уровень значимости и критическая область; - ошибки I и II рода; - связь с доверительными интервалами; - параметрические критерии; - непараметрические критерии;

№ п/п	Тематика лекционных занятий / краткое содержание
	- множественная проверка гипотез: методы контроля FWER и FDR
3	Корреляционный анализ Рассматриваемые вопросы: <ul style="list-style-type: none"> - виды связей и типы переменных; - основные виды корреляций; - корреляция Пирсона; - корреляция Спирмена; - корреляция Кенделла; - корреляция Мэтьюса; - корреляция Крамера; - бисериальный коэффициент корреляции.
4	Линейная модели Рассматриваемые вопросы: <ul style="list-style-type: none"> - основные подходы линейного анализа: регрессия и классификация; - функционалы качества; - обработка выбросов и пропусков в данных; - регуляризация; - работа с несбалансированными классами; - практическое применение.
5	Деревья решений Рассматриваемые вопросы: <ul style="list-style-type: none"> - основные принципы работы деревьев решений; - критерии разбиения: энтропийный критерий и критерий Джини; - методы поиска гиперпараметров; - практическое применение.
6	Бэггинг и бустинг Рассматриваемые вопросы: <ul style="list-style-type: none"> - принцип работы алгоритмов; - основные алгоритмы: XGBoost, LightGBM, CatBoost, Random Forest; - методы поиска гиперпараметров; - сравнение с деревом решений; - практическое применение.
7	Uplift моделирование Рассматриваемые вопросы: <ul style="list-style-type: none"> - отличие uplift моделирования от традиционных методов; - основные алгоритмы; - понятие uplift эффекта и его интерпретация; - критерии оценки качества; - визуализация uplift кривой; - практическое применение.
8	Временные ряды Рассматриваемые вопросы: <ul style="list-style-type: none"> - общие сведения о временных рядах и задачах анализа; - анализ временных рядов; - прогнозирование временных рядов; - типы временных рядов; - практическое применение.

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Доверительные интервалы В результате выполнения практического задания студент получает навыки построения доверительных интервалов и оценки точности и достоверности оценок.
2	Проверка гипотез В результате выполнения практического задания студент получает навыки формулирования гипотез, интерпретации результатов статистического теста и умение выбрать подходящий критерий для проверки гипотезы.
3	Корреляционный анализ: В результате выполнения практического задания студент получает навыки оценивания степени связи между переменными и применимости корреляционного анализа в различных ситуациях
4	Линейные модели В результате выполнения практического задания студент получает навыки понимания основных принципов линейных моделей и их применимости к различным типам данных.
5	Деревья решений В результате выполнения практического задания студент получает навыки построения и обучения деревьев решений для задач классификации и регрессии.
6	Бэггинг и бустинг В результате выполнения практического задания студент получает навыки построения и обучения случайного леса и градиентного бустинга для задач классификации и регрессии.
7	Uplift моделирование В результате выполнения практического задания студент получает навыки построения uplift моделей и их применимости в различных областях.
8	Временные ряды В результате выполнения практического задания студент получает навыки построения временных рядов и их применимости в различных областях.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение учебной литературы из приведенных источников
2	Подготовка к практическим занятиям.
3	Самостоятельное изучение лекционного материала.
4	Подготовка к промежуточной аттестации.
5	Подготовка к текущему контролю.

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

№ п/п	Библиографическое описание	Место доступа
1	G. Casella, R.L.	https://faculty.ksu.edu.sa/sites/default/files/casella_berger_statistical_inference1.pdf

	Berger. Statistical Inference – Cengage Learning, 2008 – 700 с.	
2	L. Wasserman. All of Statistics: A Concise Course in Statistical Inference –New York: Springer, 2004 – 446 с.	https://egrcc.github.io/docs/math/all-of-statistics.pdf
3	Кельберт М.Я., Сухов Ю.М. Основные понятия теории вероятностей и математической статистики – Москва: МЦНМО, 2007 – 456 с.	https://djvu.online/file/MxVgEXrKO8jR1
4	Т. Андерсон. Статистический анализ временных рядов – Москва: Мир, 2012 - 751 с. ISBN 978-5-458-26140-1	https://libcats.org/book/719600

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

- Научно-техническая библиотека РУТ (МИИТ) (<http://library.miit.ru>);
- Информационный портал Научная электронная библиотека eLIBRARY.RU (www.elibrary.ru);
- Образовательная платформа «Юрайт» (<https://urait.ru/>);
- Электронно-библиотечная система издательства «Лань» (<http://e.lanbook.com/>);
- Электронно-библиотечная система ibooks.ru (<http://ibooks.ru/>).
- Интернет-университет информационных технологий (<http://www.intuit.ru/>).

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

- Операционная система Windows;
- Microsoft Office;
- MS Teams;
- Поисковые системы.

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.

9. Форма промежуточной аттестации:

Зачет в 7 семестре.

10. Оценочные материалы.

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

Авторы:

профессор, профессор, д.н. кафедры
«Цифровые технологии управления
транспортными процессами»

А.М. Филимонов

Согласовано:

Заведующий кафедрой ЦТУТП

В.Е. Нутович

Председатель учебно-методической
комиссии

Н.А. Андриянова