

Федеральное государственное бюджетное образовательное учреждение
высшего образования
"Дальневосточный государственный университет путей сообщения"
(ДВГУПС)

УТВЕРЖДАЮ

Зав.кафедрой

(к202) Информационные технологии и
системы

Попов М.А., канд. техн.
наук, доцент



11.06.2021

РАБОЧАЯ ПРОГРАММА

дисциплины **Регрессионный и кластерный анализ**

09.03.02 Информационные системы и технологии

Составитель(и): к.т.н., Доцент, Попов М.А.

Обсуждена на заседании кафедры: (к202) Информационные технологии и системы

Протокол от 09.06.2021г. № 6

Обсуждена на заседании методической комиссии учебно-структурного подразделения: Протокол от 11.06.2021 г. № 6

Визирование РПД для исполнения в очередном учебном году

Председатель МК РНС

__ _____ 2023 г.

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2023-2024 учебном году на заседании кафедры (к202) Информационные технологии и системы

Протокол от __ _____ 2023 г. № __
Зав. кафедрой Попов М.А., канд. техн. наук, доцент

Визирование РПД для исполнения в очередном учебном году

Председатель МК РНС

__ _____ 2024 г.

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2024-2025 учебном году на заседании кафедры (к202) Информационные технологии и системы

Протокол от __ _____ 2024 г. № __
Зав. кафедрой Попов М.А., канд. техн. наук, доцент

Визирование РПД для исполнения в очередном учебном году

Председатель МК РНС

__ _____ 2025 г.

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2025-2026 учебном году на заседании кафедры (к202) Информационные технологии и системы

Протокол от __ _____ 2025 г. № __
Зав. кафедрой Попов М.А., канд. техн. наук, доцент

Визирование РПД для исполнения в очередном учебном году

Председатель МК РНС

__ _____ 2026 г.

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2026-2027 учебном году на заседании кафедры (к202) Информационные технологии и системы

Протокол от __ _____ 2026 г. № __
Зав. кафедрой Попов М.А., канд. техн. наук, доцент

Рабочая программа дисциплины Регрессионный и кластерный анализ

разработана в соответствии с ФГОС, утвержденным приказом Министерства образования и науки Российской Федерации от 19.09.2017 № 926

Квалификация **бакалавр**

Форма обучения **очная**

ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Общая трудоемкость **6 ЗЕТ**

Часов по учебному плану	216	Виды контроля в семестрах:
в том числе:		экзамены (семестр) 6
контактная работа	68	зачёты (семестр) 5
самостоятельная работа	112	РГР 6 сем. (1)
часов на контроль	36	

Распределение часов дисциплины по семестрам (курсам)

Семестр (<Курс>.<Семес тр на курсе>)	5 (3.1)		6 (3.2)		Итого	
	Неделя		16 5/6			
Вид занятий	УП	РП	УП	РП	УП	РП
Лекции	16	16	16	16	32	32
Лабораторные	16	16			16	16
Практические			16	16	16	16
Контроль самостоятельной работы			4	4	4	4
В том числе инт.	4	4	6	6	10	10
Итого ауд.	32	32	32	32	64	64
Контактная работа	32	32	36	36	68	68
Сам. работа	76	76	36	36	112	112
Часы на контроль			36	36	36	36
Итого	108	108	108	108	216	216

1. АННОТАЦИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1	Введение. Обзор методов анализа статистических данных. Дисперсионный анализ. Корреляционные зависимости. Понятие о корреляционном анализе. Регрессионный анализ. Задачи регрессионного анализа. Парная регрессионная модель. Основные предпосылки регрессионного анализа. Оценка функции регрессии. Интервальная оценка параметров парной модели. Критерий Фишера-Снедекора. Нелинейная
1.2	регрессия. Ковариационная матрица. Доверительный интервал для коэффициентов и функции регрессии. Проверка значимости уравнения регрессии. Кластерный анализ. Постановка задачи кластеризации. Алгоритм кластеризации. Гипотеза «компактности». Типы входных данных. Методы объединения кластеров. Метрики для количественных и качественных шкал. Типы кластеров: типа сгущения, с центром и

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Код дисциплины:	Б1.В.ДВ.03.02
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Научно-исследовательская работа
2.1.2	Теория вероятностей и математическая статистика
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Преддипломная практика

3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

УК-1: Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач
Знать:
Методики поиска, сбора и обработки информации; актуальные российские и зарубежные источники информации в сфере профессиональной деятельности; метод системного анализа.
Уметь:
Применять методики поиска, сбора и обработки информации; осуществлять критический анализ и синтез информации, полученной из разных источников; применять системный подход для решения составленных задач.
Владеть:
Методами поиска, сбора и обработки, критического анализа и синтеза информации; методикой системного подхода для решения поставленных задач.

ПК-1: Способность проводить исследования на всех этапах жизненного цикла программных средств

Знать:
Этапы жизненного цикла ПС и методы исследования
Уметь:
Применять методы исследования
Владеть:
Навыками исследования программных средств на всех этапах жизненного цикла

ПК-3: Способность оценивать качество программного обеспечения, в том числе проведение тестирования и исследование результатов

Знать:
Методы оценки качества ПО, тестирования и исследование результатов
Уметь:
Проводить тестирование и методы обработки результатов
Владеть:
Навыками обоснования правильности выбранного метода оценки качества ПО и обработкой результатов исследования.

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ), СТРУКТУРИРОВАННОЕ ПО ТЕМАМ (РАЗДЕЛАМ) С УКАЗАНИЕМ ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература	Инте ракт.	Примечание
	Раздел 1. Методы анализа статистических данных						

1.1	Анализ данных. Основные понятия и задачи анализа данных. Виды данных. Методы анализа данных. Дескриптивная статистика: визуализация распределения единиц совокупности, вычисление мер центральной тенденции, оценка разброса данных, выявление выбросов. /Лек/	5	2	УК-1 ПК-1 ПК-3	Л1.7Л2.1 Л2.2 Л2.3	0	
1.2	Проверка гипотез о параметрах генеральной совокупности: о законе распределения вероятностей случайной величины (критерий Хи-квадрат Пирсона), о равенстве математических ожиданий нормальных генеральных совокупностей (критерий Стьюдента), о равенстве дисперсий нормальных генеральных совокупностей (критерий Фишера) /Лек/	5	2	УК-1 ПК-1 ПК-3	Л1.1 Л1.6 Л1.5 Л1.4Л2.1 Л2.2 Л2.3 Л2.4	0	
Раздел 2. Корреляционно-регрессионный анализ							
2.1	Понятие о корреляционном анализе. Корреляционные зависимости. /Лек/	5	2	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.7 Л1.6 Л1.3Л2.1 Л2.2 Л2.3	0	
2.2	Регрессионный анализ. Задачи регрессионного анализа. Парная регрессионная модель. Основные предпосылки регрессионного анализа. Оценка функции регрессии. Интервальная оценка параметров парной модели. Критерий Фишера-Снедекора. /Лек/	5	4	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.7 Л1.6 Л1.3Л2.1 Л2.2 Л2.3	0	
2.3	Нелинейная регрессия. Ковариационная матрица. Доверительный интервал для коэффициентов и функции регрессии. Проверка значимости уравнения регрессии. /Лек/	5	4	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.7 Л1.6 Л1.3Л2.1 Л2.2 Л2.3	4	Активное слушание
2.4	Определение коэффициента корреляционной зависимости /Лаб/	5	4	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.2 Л1.7 Л1.6Л2.1 Л2.2 Л2.3 Э2 Э3	0	
2.5	Построение парной линейной регрессионной модели и ее анализ /Лаб/	5	4	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.2 Л1.7 Л1.6Л2.1 Л2.2 Л2.3 Э2	0	
2.6	Построение парной нелинейной регрессионной модели и ее анализ /Лаб/	5	4	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.2 Л1.7 Л1.6 Л1.3Л2.1 Л2.2 Л2.3 Э2 Э3	0	
Раздел 3. Дисперсионный анализ							
3.1	Дисперсионный анализ. Парная и множественная модель /Лек/	5	2	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.7 Л1.6 Л1.3Л2.2 Л2.3	0	

3.2	Дисперсионный анализ данных /Лаб/	5	4	УК-1 ПК-1 ПК-3	Л1.1 Л1.7Л2.2 Л2.3 Э1 Э2	0	
Раздел 4. Самостоятельная работа							
4.1	Подготовка к аудиторным занятиям, работа с литературой, подготовка отчетов по лабораторным работам /Ср/	5	68	УК-1 ПК-1 ПК-3	Л1.1 Л1.7Л2.2 Л2.3	0	
4.2	Подготовка к зачету /Ср/	5	8	УК-1 ПК-1 ПК-3	Л1.7Л2.2	0	
Раздел 5. Кластерный анализ							
5.1	Кластерный анализ. Постановка задачи кластеризации. Алгоритм кластеризации. Гипотеза компактности». Типы входных данных. Методы объединения кластеров. Метрики для количественных и качественных шкал. Типы кластеров: типа сгущения, с центром и другие. /Лек/	6	4	УК-1 ПК-1 ПК-3	Л1.2 Л1.7Л2.2 Л2.4	0	
5.2	Алгоритмы и методы кластеризации данных /Пр/	6	4	УК-1 ПК-1 ПК-3	Л1.7Л2.2 Л2.4 Э3 Э4	0	
Раздел 6. Непараметрические методы анализа данных							
6.1	Задачи, решаемые с помощью непараметрических методов анализа данных, условия их применения. Оценка взаимосвязи на основе таблиц сопряженности. Показатели связи в случае дихотомических и многовариантных переменных. Измерение связи с помощью методов ранговой корреляции. Понятие ранга. Связанные ранги. Коэффициенты ранговой корреляции. Оценка различий между независимыми группами. Критерий суммы рангов Вилкоксона для сравнения двух независимых выборок. Статистика Манна-Уитни. Критерий Краскала-Уоллиса для сравнения двух и более независимых выборок. /Лек/	6	4	УК-1 ПК-1 ПК-3	Л1.2 Л1.7Л2.2 Л2.4	0	
6.2	Оценка взаимосвязи на основе таблиц сопряженности. Измерение связи с помощью методов ранговой корреляции. Коэффициенты ранговой корреляции. Применение статистик для оценок /Пр/	6	4	УК-1 ПК-1 ПК-3	Л1.2 Л1.7Л2.2 Л2.4	0	
Раздел 7. Дискриминантный анализ							
7.1	Каноническая дискриминация. Коэффициенты канонической дискриминантной функции. Нестандартизованные коэффициенты. Число дискриминантных функций /Лек/	6	4	УК-1 ПК-1 ПК-3	Л1.7Л2.2 Л2.4	2	Активное слушание
7.2	Классифицирующие функции. Применение элементарных классифицирующих функций. Классификация объектов с помощью функции расстояния. Классификационная матрица /Лек/	6	4	УК-1 ПК-1 ПК-3	Л1.7Л2.2 Л2.4	0	

7.3	Каноническая дискриминация. /Пр/	6	4	УК-1 ПК-1 ПК-3	Л1.2 Л1.7Л2.2 Л2.4 Э2 Э3	4	Работа в малых граппах
7.4	Классификационные функции /Пр/	6	4	УК-1 ПК-1 ПК-3	Л1.7Л2.2 Л2.4 Э2	0	
Раздел 8. Самостоятельная работа							
8.1	Выполнение РГР /Ср/	6	8	УК-1 ПК-1 ПК-3	Л1.7Л2.2	0	
8.2	Подготовка к аудиторным занятиям, работа с литературой, подготовка отчетов по лабораторным работам /Ср/	6	28	УК-1 ПК-1 ПК-3	Л1.7Л2.2	0	
Раздел 9. Контроль							
9.1	Подготовка к экзамену /Экзамен/	6	36	УК-1 ПК-1 ПК-3	Л1.8 Л1.1 Л1.2 Л1.7Л2.1 Л2.2 Л2.3 Л2.4 Э1 Э2 Э3 Э4	0	

5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Размещены в приложении

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Перечень основной литературы, необходимой для освоения дисциплины (модуля)

	Авторы, составители	Заглавие	Издательство, год
Л1.1	Чашкин Ю.Р.	Прикладная статистика. Статистическая обработка данных: Метод. указания	Хабаровск: Изд-во ДВГУПС, 2005,
Л1.2	Бондаренко С.В., Бондаренко М.Ю.	Excel 2007. Популярный самоучитель	Санкт-Петербург: Питер, 2008,
Л1.3	Постовалов С. Н., Лемешко Б. Ю., Лемешко С. Б., Чимитова Е. В.	Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход	Москва: ООО "Научно-издательский центр ИНФРА-М", 2015, http://znanium.com/go.php?id=515227
Л1.4	Жуковский О. И.	Информационные технологии и анализ данных: учебное пособие	Томск: Эль Контент, 2014, http://biblioclub.ru/index.php?page=book&id=480500
Л1.5	Крутиков В. Н., Мешечкин В. В.	Анализ данных: учебное пособие	Кемерово: Кемеровский государственный университет, 2014, http://biblioclub.ru/index.php?page=book&id=278426
Л1.6	Воскобойников Ю.Е.	Регрессионный анализ данных в пакете Mathcad: учеб. пособие	Москва: Лань, 2011, http://e.lanbook.com/books/element.php?pl1_cid=25&pl1_id=666
Л1.7	Крутиков В. Н., Мешечкин В. В.	Анализ данных	Кемерово: Кемеровский государственный университет, 2014, http://biblioclub.ru/index.php?page=book&id=278426
Л1.8	Тюрин Ю.Н., Макаров А.А.	Статистический анализ данных на компьютере	Москва: ИНФРА-М, 1998,

6.1.2. Перечень дополнительной литературы, необходимой для освоения дисциплины (модуля)

	Авторы, составители	Заглавие	Издательство, год
Л2.1	Дворянкина Е.К., Коровина С.В.	Математические методы обработки данных: метод. пособие	Хабаровск: Изд-во ДВГУПС, 2013,

	Авторы, составители	Заглавие	Издательство, год
Л2.2	Орлов А. И.	Прикладная статистика	Москва: Интернет- Университет Информационных Технологий, 2009, http://biblioclub.ru/index.php?page=book&id=234537
Л2.3	Ю.В. Пономарчук, А.И. Кондратьев	Прикладная статистика Ч.1: учеб.пособие. В 2 ч.	Хабаровск: Изд-во ДВГУПС, 2007,
Л2.4	Ю.В. Пономарчук, А.И. Кондратьев	Прикладная статистика Ч.2 : учеб.пособие. В 2 ч.	Хабаровск: Изд-во ДВГУПС, 2007,

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Э1	Loginorm	https://loginom.ru/for/analyst?utm_source=direct&utm_campaign=62921554&utm_content=gid 4694135613 ad_id 11164329385 phrase_id 34046022569&utm_term=%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0%20%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0&type=search&added=no&block=premium&pos=1&device=desktop&yclid=18232811058497381457
Э2	Statsoft	http://statsoft.ru/home/textbook/modules/stdiscan.html
Э3	Анализ больших данных в EXCEL	https://1-a-b-a.com/blog/show/158
Э4	Анализ данных	http://repo.ssau.ru/bitstream/Uchebnye-posobiya/Analiz-dannyh-Elektronnyi-resurs-ucheb-posobie-po-programme-vysshe-obrazovaniya-59138/1/%D0%9F%D0%BE%D1%80%D1%83%D1%87%D0%B8%D0%BA%D0%BE%D0%B2%20%D0%9C.%D0%90.%20%D0%90%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7%20%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85.pdf

6.3 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

6.3.1 Перечень программного обеспечения

Free Conference Call (свободная лицензия)

Zoom (свободная лицензия)

Matlab Базовая конфигурация (Academic new Product Concurrent License в составе: (Matlab, Simulink, Partial Differential Equation Toolbox) - Математический пакет, контракт 410

Office Pro Plus 2007 - Пакет офисных программ, лиц.45525415

6.3.2 Перечень информационных справочных систем

Профессиональная база данных, информационно-справочная система Гарант - <http://www.garant.ru>

Профессиональная база данных, информационно-справочная система КонсультантПлюс - <http://www.consultant.ru>

7. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Аудитория	Назначение	Оснащение
201	Компьютерный класс для практических и	столы, стулья, компьютерная техника с возможностью

Аудитория	Назначение	Оснащение
	лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также для самостоятельной работы	к сети Интернет, свободному доступу в ЭБС и ЭИОС, проектор
304	Учебная аудитория для проведения занятий лекционного типа	комплект учебной мебели: столы, стулья, интерактивная доска, мультимедийный проектор, компьютер, система акустическая
424	Учебная аудитория для проведения лабораторных и практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Лаборатория электронных устройств регистрации и передачи информации	комплект учебной мебели, мультимедийный проектор, экран, компьютер преподавателя

8. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Видами аудиторной работы студента при изучении дисциплины являются лекции, лабораторные и практические занятия. Студент обязан посещать аудиторные занятия. На лекциях излагаются и разъясняются основные понятия темы, связанные с ней теоретические и практические проблемы, даются рекомендации для самостоятельной работы. В ходе лекции студент должен внимательно слушать и конспектировать лекционный материал. При необходимости студент имеет право задать вопрос в отношении изложенного материала во время, отведенное для этих целей преподавателем.

На практических занятиях излагаются и разъясняются основные понятия темы, связанные с выполнением практических заданий, даются рекомендации для самостоятельной работы и выполнения РГР. При подготовке к практическим занятиям студент должен изучить вопросы ранее рассмотренные на лекционных занятиях.

По организации самостоятельной работы

Для студентов самостоятельная работа является одним из основных видов работы по изучению дисциплины. Она включает

- изучение материала установочных занятий;
- работу с рекомендованной литературой и дополнительными источниками информации;
- подготовку к сдаче зачета и экзамена.

Самостоятельную работу по изучению дисциплины целесообразно начинать с изучения рабочей программы, которая содержит основные требования к знаниям, умениям, навыкам обучаемых, ознакомления с разделами и темами.

Для лиц с ограниченными возможностями используются дистанционные образовательные технологии, а именно сайт ДВГУПС <http://www.dvgups.ru/> и рабочая программа дисциплины.