



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МСХА имени К.А. ТИМИРЯЗЕВА»
(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)

ФЕДЕРАЛЬНЫЙ ЦЕНТР КОМПЕТЕНЦИЙ



Проректор по учебной работе
Е.В. Хохлова
«06» апреля 2023г.

ПРОГРАММА ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

СТАТИСТИЧЕСКИЙ АНАЛИЗ СЕЛЬСКОГО ХОЗЯЙСТВА
С ИСПОЛЬЗОВАНИЕМ ЯЗЫКА ПРОГРАММИРОВАНИЯ R

г. Москва, 2023

РАЗДЕЛ 1. ХАРАКТЕРИСТИКА ПРОГРАММЫ

1.1. Цель реализации программы

Совершенствование и/или приобретение новых профессиональных компетенций слушателями в области статистического анализа сельского хозяйства с использованием языка программирования R.

Совершенствуемые и/или приобретаемые компетенции и планируемые результаты обучения

№	Приобретаемые и/или совершенствуемые компетенции	Код компетенции	Знать/Уметь
1.	Способен применять естественнонаучные и общетехнические знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности;	ОПК-1	Знать: основы математики, вычислительной техники и программирования Уметь: решать стандартные профессиональные задачи с применением естественнонаучных и общетехнических знаний, методов математического анализа и моделирования
2.	Способность проводить анализ данных с использованием информационных технологий в области сельского хозяйства, экономики, бухгалтерского учета, статистики, финансов и др.	ПКос-8	Знать: информационные технологии анализа данных Уметь: собирать информацию для проведения анализа; устанавливать причинно-следственные связи между признаками; выбирать и применять, в том числе с использованием современных информационных технологий, методы анализа данных; делать выводы на основе проведенного анализа данных
3.	Способен проводить аналитическое исследование с применением технологий больших данных	ПКос-9	Знать: теоретические и прикладные основы анализа, современные методы и инструментальные средства анализа данных Уметь: автоматизировать процесс анализа данных; визуализировать результаты анализа данных

РАЗДЕЛ 2. СОДЕРЖАНИЕ ПРОГРАММЫ

2.1. Учебный план программы повышения квалификации «Статистический анализ сельского хозяйства с использованием языка программирования R»

Категория слушателей: руководители и специалисты области АПК, преподаватели образовательных учреждений, сотрудники НИИ, аспиранты, докторанты, студенты ВУЗов и все, кто интересуется вопросами статистического анализа сельского хозяйства с использованием языка программирования R.

Форма обучения : дистанционная.

Срок освоения: 12 недель

Трудоемкость программы: 36 академических часов

№ п/п	Наименование разделов, тем	Всего ак. часов	В том числе			Формы аттестации, контроля
			Сам. работа	Лекции	Практ. занятия	
1	Раздел 1 Основы работы с языком программирования R	8	2	2	4	Практическая работа
2	Раздел 2 Статистический анализ сельского хозяйства в среде R	14	4	4	6	Практическая работа
3	Раздел 3 Интеллектуальный анализ данных в среде R	14	6	4	4	Практическая работа
Итоговая аттестация		Зачет				

2.2. Учебно-тематический план программы повышения квалификации «Статистический анализ сельского хозяйства с использованием языка программирования R»

№ п/п	№ раздела	Виды учебных занятий, кол-во ак. часов	Содержание	Планируемый результат
1	2	3	4	5
1	Раздел I Основы работы с языком программи	Лекция 1, 2 ак.ч.	Основы работы в среде R	Знать основы программирования в среде R
		Практическая работа №1, 4 ак.ч.	Основы программирования на языке R	Уметь решать стандартные профессиональные

№ п/п	№ раздела	Виды учебных занятий, кол-во ак. часов	Содержание	Планируемый результат
1	2	3	4	5
	рования R			задачи с применением естественнонаучных и общетехнических знаний, методов математического анализа и моделирования
		Самостоятельная работа, 2 ак.ч.	Разработка собственных функций в среде R	Знать и уметь использовать язык программирования R для автоматизации процесса анализа данных и визуализации результатов
2	Раздел II Статистический анализ сельского хозяйства в среде R	Лекция 2, 1 ак.ч.	Ряды распределения и их количественные характеристики	Знать основы применения математического аппарата в статистике
		Лекция 3, 2 ак.ч.	Корреляционно-регрессионный анализ данных	Знать теоретические и прикладные основы анализа, методы корреляционно-регрессионного анализа
		Лекция 4, 1 ак.ч.	Ряды динамики	Знать теоретические и прикладные основы анализа, современные методы и инструментальные средства анализа данных временных рядов
		Практическая работа №2, 2 ак.ч.	Основные статистики и проверка статистических гипотез в среде R	Уметь устанавливать причинно-следственные связи между признаками; выбирать и применять методы анализа данных в среде R; делать выводы на основе проведенного анализа данных
		Практическая работа №3, 2 ак.ч.	Корреляционно-регрессионный анализ в среде R	Уметь устанавливать причинно-следственные связи между признаками; выбирать и применять корреляционно-регрессионный анализ в среде R; делать выводы на основе проведенного анализа данных
		Практическая работа №4,	Ряды динамики в	Уметь устанавливать

№ п/п	№ раздела	Виды учебных занятий, кол-во ак. часов	Содержание	Планируемый результат
1	2	3	4	5
		2 ак.ч.	среде R	причинно-следственные связи между признаками; выбирать и применять методы анализа рядов динамики в среде R; делать выводы на основе проведенного анализа данных
		Самостоятельная работа, 4 ак.ч.	Проведение анализа данных в среде R	Уметь устанавливать причинно-следственные связи между признаками; выбирать и применять методы анализа данных в среде R; делать выводы на основе проведенного анализа данных
3	Раздел III Ряды динамики в среде R	Лекция 5, 2 ак.ч.	Введение в интеллектуальный анализ данных	Знать основы интеллектуального анализа данных и информационные технологии анализа данных
		Лекция 6, 2 ак.ч.	Методы интеллектуального анализа данных	Знать методы интеллектуального анализа данных
		Практическая работа №5, 2 ак.ч.	Построение метода «Дерево решений» в среде R	Уметь автоматизировать процесс интеллектуального анализа данных; визуализировать результаты анализа данных, делать выводы на основе проведенного анализа данных
		Практическая работа №6, 2 ак.ч.	Алгоритм «Случайный лес» в среде R	Уметь автоматизировать процесс интеллектуального анализа данных, делать выводы на основе проведенного анализа данных
		Самостоятельная работа, 6 ак.ч.	Разработка моделей и настройка параметров методами «Дерево решений» и «Случайный лес»	Уметь автоматизировать процесс интеллектуального анализа данных, делать выводы на основе проведенного анализа данных

Раздел 3. Формы аттестации и оценочные материалы

Входное тестирование

Форма проведения	Очно
Виды оценочных материалов	Тест из 10 заданий в электронной форме (Приложение 1)
Критерии оценивания	1 – правильный ответ; 0 – неправильный ответ. 8-10 баллов – высокий уровень, 5-7 баллов – средний уровень, менее 5 – низкий уровень.
Оценка	Не предусмотрено (тестирование проводится с целью определения уровня владения материалом)

Выходное тестирование

Форма проведения	Очно
Виды оценочных материалов	Тест из 20 заданий в электронной форме (Приложение 3)
Критерии оценивания	1 – правильный ответ; 0 – неправильный ответ. «Зачтено» выставляется слушателям, если они набрали 11 баллов
Оценка	Зачтено/не зачтено

Практическая работа 1 к разделу 1

Название	Основы программирования на языке R
Структура и содержание	Изучить основы работы с языком программирования R Структура работы: 1. Изучить основные объекты языка; 2. Основные функции работы с векторами; 3. Работа с матрицами и Data Frame; 4. Построение графиков в среде R; 5. Создание и использование функций на языке R.
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности действий.
Оценка	Зачтено/не зачтено

Практическая работа 2 к разделу 2

Название (проекта, разработки, сценария и т.д.)	Основные статистики и проверка статистических гипотез в среде R
Структура и содержание	Для рассматриваемого набора данных определить соответствие исходных показателей нормальному закону распределения и рассчитать основные показатели вариации признаков. Структура работы: 1. Загрузить исходные данные в R; 2. Рассчитать основные статистики; 3. Построить гистограмму распределений; 4. Проверить гипотезу о соответствии данных нормальному закону распределения.
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности действий.

Оценка	Зачтено/не зачтено
--------	--------------------

Практическая работа 3 к разделу 2

Название	Корреляционно-регрессионный анализ в среде R
Структура и содержание	Для рассматриваемых показателей, характеризующих сельскохозяйственные предприятия, построить модель корреляционно-регрессионного анализа и описать результаты. Структура работы: 1. Построить диаграммы рассеяния и матрицу парных коэффициентов корреляции; 2. построить модель парной линейной регрессии, оценить тесноту связи в уравнении регрессии, оценить достоверность уравнения в целом с использованием критерия F-Фишера, оценить достоверность отдельных параметров уравнения, дать интерпретацию коэффициента полной регрессии, проверить модель на соответствие предпосылкам метода наименьших квадратов; 3. построить множественную модель регрессии используя пошаговую процедуру включения-исключения факторов, оценить качество построенного уравнения;
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности действий.
Оценка	Зачтено/не зачтено

Практическая работа 4 к разделу 2

Название	Ряды динамики в среде R
Структура и содержание	Требуется с помощью языка программирования R проанализировать динамику исходного ряда, проверить ряд на наличие структурных сдвигов и построить модели для прогнозирования с помощью линейной регрессии и экспоненциального сглаживания. Структура работы: 1. Загрузить исходные данные и преобразовать их в ряд динамики; 2. Построить график динамического ряда; 3. Проверить исходный ряд на наличие структурных сдвигов; 4. построить модели для прогнозирования с помощью линейной регрессии и экспоненциального сглаживания; 5. Сравнить качество построенных моделей.
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности действий.
Оценка	Зачтено/не зачтено

Практическая работа 5 к разделу 3

Название	Построение метода «Дерево решений» в среде R
Структура и содержание	Разработать модели классификации методом «Дерево решений» с обучением и без. Настроить параметры алгоритмов. Оценить качество классификации и выбрать наилучший алгоритм. Структура работы: 1. Построить модель классификации методом дерева решений с обучением и без; 2. Настроить параметры алгоритма дерева решений; 3. Сравнить метрики качества построенных моделей;
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в

	полном объеме с соблюдением необходимой последовательности действий.
Оценка	Зачтено/не зачтено

Практическая работа 6 к разделу 3

Название	Алгоритм «Случайный лес» в среде R
Структура и содержание	Разработать модели классификации методом «Случайный лес» с обучением и без. Настроить параметры алгоритмов. Оценить качество классификации и выбрать наилучший алгоритм. Структура работы: 1. Построить модель классификации методом «Случайный лес» с обучением и без; 2. Настроить параметры алгоритма «Случайный лес»; 3. Сравнить метрики качества построенных моделей;
Критерии оценивания	Оценка «зачтено» ставится, если обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности действий.
Оценка	Зачтено/не зачтено

Итоговая аттестация

Форма итоговой аттестации	Зачет как совокупность выполненного итогового теста и практических работ
Требования к итоговой аттестации	Выполнение итогового теста и практических работ в соответствии с требованиями к каждой из работ
Критерии оценивания	Слушатель считается аттестованным при положительном оценивании практических работ и итогового тестирования
Оценка	Зачтено/не зачтено

Раздел 4. Материально-технические условия реализации программы

Для реализации программы используются ресурсы, размещенные в системе дистанционного обучения ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева (sdo.timacad.ru), которые позволяют слушателям самостоятельно осваивать содержание программы или отдельных ее разделов, используются MOOK, открытые образовательные и интернет – ресурсы и платформы.

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
1	2	3
Аудитория	Лекции, внеаудиторные учебные занятия	компьютеры, операционная система: <i>Windows - 64-битная x86, 32-битная x86; MacOS - 64-битная x86; Linux - 64-битная x86, 64-битная Power8 /</i>

		<i>Power, web-браузер - любой из перечисленных: Chrome, Safari, Internet Explorer, Anaconda</i>
LMS Moodle (дистанционная образовательная платформа ФГБОУ ВО РГАУ МСХА имени К.А. Тимирязева)	Практические и лабораторные занятия	Sdo.timacad.ru Доступ в сеть интернет, компьютеры и программное обеспечение, поддерживающее работу сайта

5. Учебно-методическое обеспечение программы

Основная литература:

1. Гришин, В. А. Основы программирования на языке R : учебно-методическое пособие / В. А. Гришин. — Нижний Новгород : ННГУ им. Н. И. Лобачевского, 2021. — 67 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/191498> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
2. Митина, О. А. Языки программирования для статистической обработки данных (R) : учебное пособие / О. А. Митина. — Москва : РТУ МИРЭА, 2020. — 191 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163912> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
3. Гришин, В. А. Методы обработки данных и моделирование на языке R : учебно-методическое пособие / В. А. Гришин, М. С. Тихов. — Нижний Новгород : ННГУ им. Н. И. Лобачевского, 2019. — 54 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/144653> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
4. Перекатов, А. С. Статистическая обработка экспериментальных данных. Полный факторный эксперимент в языке R : учебное пособие / А. С. Перекатов, М. Б. Никифоров. — Рязань : РГРТУ, 2019. — 48 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/168309> (дата обращения: 28.11.2022). — Режим доступа: для авториз. пользователей.
5. Data Science / Francesco Palumbo, Angela Montanari, Maurizio Vichi. Springer International Publishing AG, 2017 – Текст : электронный // Springer: электронно-библиотечная система. URL: <https://link.springer.com/book/10.1007/978-3-319-55723-6#editorsandaffiliations> (дата обращения: 08.04.2023).

Дополнительная литература:

1. Алексеев, Д. С. Технологии интеллектуального анализа данных : учебник для вузов / Д. С. Алексеев, О. В. Щекочихин. — Санкт-Петербург : Лань, 2022. — 176 с. — ISBN 978-5-8114-8299-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/187559>
2. Макшанов, А. В. Технологии интеллектуального анализа данных : учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2022. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/206711>

3. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : НГТУ, 2017. — 74 с. — ISBN 978-5-7782-3183-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/118287> (дата обращения: 15.11.2022). — Режим доступа: для авториз. Пользователей.
4. Гильванов, Р. Г. Технологии обработки информации: Лабораторный практикум : учебное пособие / Р. Г. Гильванов, , А. В. Забродин. — Санкт-Петербург : ПГУПС, 2022. — 45 с. — ISBN 978-5-7641-1819-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/279065> (дата обращения: 28.11.2022). — Режим доступа: для авториз. пользователей.

Интернет-ресурсы:

1. The R Project for Statistical Computing <https://www.r-project.org/> (открытый доступ)
2. Анаконда. URL: <https://www.anaconda.com/distribution/> (открытый доступ) (дата обращения: 08.04.2023).
2. Официальный сайт Росстата. URL: <https://rosstat.gov.ru/> (открытый доступ) (дата обращения: 08.04.2023).

6. Оценка качества освоения программы

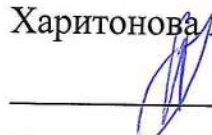
Оценка качества освоения программы осуществляется на основе результатов итоговой аттестации. Слушатель считается аттестованным, если имеет средний уровень и выше усвоения материала по всем разделам программы.

7. Образовательные технологии, используемые в процессе реализации программы

В программе используются ресурсы, размещенные в системе дистанционного обучения ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева (sdo.timacad.ru), которые позволяют слушателям самостоятельно осваивать содержание программы или отдельных ее разделов, используются МООК, открытые образовательные и интернет – ресурсы и платформы.

8. Составители программы

Харитонов А.Е., канд.экон.наук., доцент (разделы 1-3)

 / А.Е. Харитонов /

Кагирова М.В., канд.экон.наук., доцент (разделы 2)

 / М.В. Кагирова /

Токарев В.С., ассистент (разделы 1)

 / В.С. Токарев /

Утверждено на кафедре статистики и кибернетики Института экономики и управления АПК

Протокол № 08 от «06» апреля 2023 г.

И.о.зав. кафедрой статистики и кибернетики  / А.В. Уколова /

Приложение 1

Входной тест

1. Статистическая совокупность – это:

- a) множество единиц с варьирующим значением признака
- b) множество чисел, характеризующих значение признака
- c) множество градаций значений признака
- d) множество единиц с одинаковым значением признака

2. Составной элемент объекта статистического изучения, являющийся носителем признаков, подлежащих регистрации, называется

- a) единицей наблюдения
- b) единицей регистрации
- c) статистическим формуляром
- d) объектом наблюдения

3. Непрерывный признак группировок:

- a) число членов семьи
- b) пол человека
- c) заработная плата работающих
- d) национальность человека

4. К видам несплошного статистического наблюдения относится:

- a) регистровое наблюдение
- b) обследование основного массива
- c) специально организованное наблюдение
- d) текущее статистическое наблюдение

5. Что представляет собой статистическое наблюдение

- a) планомерную, научно обоснованную деятельность по сбору данных о состоянии массовых общественных явлений путем регистрации значений их признаков по единицам статистических совокупностей;
- b) метод научного познания мира, состоящего в целенаправленном восприятии действительности при помощи органов чувств
- c) статистическая методология регистрации значений признаков массовых общественных явлений по единицам статистических совокупностей.

6. При формировании официальной статистической документированной информации используются следующие виды данных:

- a) абсолютные, средние и относительные показатели

- b) первичные статистические данные и административные данные
- c) федеральные, региональные, муниципальные данные и данные предприятий
- d) экономические, социальные и экологические показатели развития общества

5. Изучение взаимосвязи между двумя переменными возможно благодаря...

- a) двумерному графику
- b) гистограммам
- c) графикам временных рядов
- d) ни один из вариантов

6. График может быть удобным представлением данных, если...

- a) существует взаимосвязь между объектами данных
- b) объекты данных показывают определенную тенденцию
- c) оба варианта верны
- d) ни один из вариантов не является верным

7. Первым этапом визуализации данных является...

- a) беглый обзор
- b) построение графика
- c) интерпретация результатов
- d) оценка эффективности

8. Что из перечисленного не является графиком?

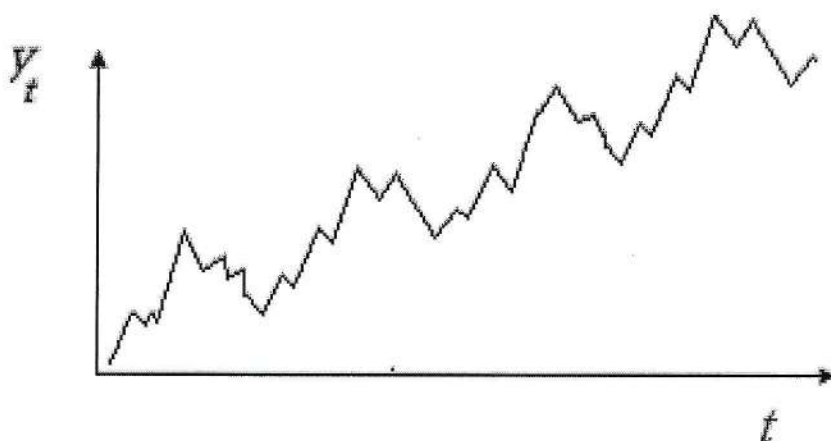
- a) гистограмма
- b) круговая диаграмма
- c) ранжированный ряд
- d) лепестковая диаграмма

9. Визуализация данных позволяет нам обнаружить...

- a) закономерности
- b) тренды
- c) корреляции
- d) всё выше перечисленное

Выходной тест

1. **Какая функция на языке программирования R используется для построения корреляционно-регрессионной модели:**
 - a) mod
 - b) lm
 - c) cor
 - d) var
2. **Функция `bptest` пакеты `lmtest` на языке программирования R используется для:**
 - a) Проведения теста на гетероскедастичность остатков
 - b) Проведения теста на автокорреляцию остатков
 - c) Проведения теста на соответствия нормальному закону распределения
 - d) Проверки модели на устойчивость
3. **Какой пакет языка программирования R используется для анализа временных рядов?**
 - a) rpart
 - b) prophet
 - c) e1071
 - d) forcats
4. **Из каких компонент состоит временной ряд, представленный на графике:**



- a) ряд содержит циклическую компоненту
 - b) ряд содержит тенденцию, случайную и циклическую компоненты
 - c) ряд содержит случайную компоненту
 - d) ряд содержит тенденцию и случайную компоненту
5. **По результатам построения парной модели регрессии на языке программирования R определите достоверность параметров уравнения для генеральной совокупности:**

```

Call:
lm(formula = y ~ x)

Residuals:
    Min       1Q   Median       3Q      Max
-1.38822 -0.29358 -0.04393  0.26429  1.34521

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.77763    0.07293   65.51  <2e-16 ***
x            0.88858    0.05137   17.30  <2e-16 ***
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.478 on 148 degrees of freedom
Multiple R-squared:  0.669,    Adjusted R-squared:  0.6668
F-statistic: 299.2 on 1 and 148 DF,  p-value: < 2.2e-16

```

- a) Достоверен только коэффициент полной регрессии
- b) Оба параметра достоверны
- b) Достоверно только условное начало
- d) Оба параметра не достоверны

6. Что такое интеллектуальный анализ данных?

- a) извлечение информации из необработанных данных
- b) хранение и обработка информации в базах данных
- c) визуализация данных
- d) статистический метод обработки данных

7. Что является результатом представления знаний в интеллектуальном анализе данных?

- a) закономерность
- b) проект
- c) модель
- d) тренд

8. Задача классификации – это _____ задача.

- a) описательная
- b) предсказательная
- c) качественная
- d) количественная

9. Задача кластеризации – это _____ задача.

- a) описательная
- b) предсказательная
- c) качественная
- d) количественная

10. Единицы наблюдения, которые значительно отличаются от большинства других единиц в наборе данных:

- a) транзакция
- b) порядковое число
- c) интервалы

d) резко выделяющиеся значения

11. Какая иерархическая структура у деревьев принятия решений?

- a) ЕСЛИ... ТО...
- b) НИ... НИ...
- c) ЛИБО... ЛИБО...
- d) КАК... ТАК И...

12. Что из перечисленного является математическим уравнением, связывающим переменные x и y ?

- a) регрессия
- b) интерполяция
- c) кластеризация
- d) экстраполяция

13. Уравнение вида $y = a + bx$:

- a) полиномиальное уравнение
- b) линейная регрессия
- c) регрессия
- d) интерполяция

14. Какой из этих показателей является измерением качества модели регрессии?

- a) средняя арифметическая
- b) дисперсия
- c) стандартное отклонение
- d) коэффициент детерминации

15. Коэффициент корреляции всегда лежит между значениями...

- a) 0 и 1
- b) -1 и 1
- c) -1 и 0
- d) 0 и 2

16. Изучение взаимосвязи между несколькими переменными – это задача...

- a) парной регрессии
- b) множественной регрессии
- c) дерева принятия решения
- d) моделирования

17. Ожидаемое значение y , когда $X = 0$, равно:

- a) коэффициенту полной регрессии
- b) условному началу
- c) коэффициенту корреляции
- d) коэффициенту детерминации

18. Собранные в разные моменты времени значения каких-либо параметров - это...

- a) панельные данные
- b) пространственные данные
- c) данные временного ряда
- d) ни один из вариантов

19. Укажите условия, которые соответствуют требованиям МНК:

- a) неслучайный характер остатков
- b) остатки подчиняются нормальному закону распределения
- c) автокорреляция остатков
- d) гетероскедастичность остатков

20. Выберите правильный код на языке R для построения модели случайный лес в пакете «randomForest»:

- a)

```
train<-dt[-seq(1, nrow(dt), 3),]
test<-dt[seq(1, nrow(dt), 3),]
RF<-randomForest(train$y~., data=train)
pred<-predict(RF, test)
table(pred, test$y)
```
- b)

```
dt_mixed<-dt[order(runif(150)),]
train<-dt_mixed[1:100,]
test<-dt_mixed[101:150,]
RF<-randomForest(dt$y~., data=dt)
pred<-predict(RF, test)
table(pred, test$y)
```
- c)

```
train<-dt[-seq(1, nrow(dt), 3),]
test<-dt[seq(1, nrow(dt), 3),]
RF<-randomForest(test$y~., data=test)
pred<-predict(RF, train)
table(pred, train$y)
```
- d)

```
dt_mixed<-dt[order(runif(150)),]
train<-dt_mixed[1:100,]
test<-dt_mixed[101:150,]
RF<-randomForest(dt$y~., data=dt)
pred<-predict(RF, train)
table(pred, train $y)
```


**Профессиональные компетенции по направлению подготовки
09.03.02 Информационные системы и технологии**

ОПК-1	Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности;
ОПК-2	Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности
ОПК-3	Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности
ОПК-4	Способен участвовать в разработке технической документации, связанной с профессиональной деятельностью с использованием стандартов, норм и правил
ОПК-5	Способен устанавливать программное и аппаратное обеспечение для информационных и автоматизированных систем;
ОПК-6	Способен разрабатывать алгоритмы и программы, пригодные для практического применения в области информационных систем и технологий
ОПК-7	Способен осуществлять выбор платформ и инструментальных программно-аппаратных средств для реализации информационных систем
ОПК-8	Способен применять математические модели, методы и средства проектирования информационных и автоматизированных систем.
ПКос-1	Способность проводить исследования в профессиональной деятельности, составлять отчеты о проделанной работе, готовить публикации, выступать с докладами
ПКос-2	Способность выполнять интеграцию программных модулей и компонент
ПКос-3	Способность выполнять работы по обеспечению функционирования баз данных и обеспечению их информационной безопасности
ПКос-4	Способность выполнять работы по созданию (модификации) и сопровождению информационных систем
ПКос-5	Способность создания технической документации на продукцию в сфере информационных технологий, управления технической информацией
ПКос-6	Способность выполнять работы по обслуживанию программно-аппаратными средствами сетей и инфокоммуникаций
ПКос-7	Способность использовать компоненты системных программных продуктов: компиляторы, загрузчики, сборщики и системные утилиты
ПКос-8	Способность проводить анализ данных с использованием информационных технологий в области сельского хозяйства, экономики, бухгалтерского учета, статистики, финансов и др.
ПКос-9	Способен проводить аналитическое исследование с применением технологий больших данных