

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.О.31 Анализ данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

01.03.04 Прикладная математика

Направленность (профиль)

01.03.04 Прикладная математика

Форма обучения

очная

Год набора

2022

Красноярск 2023

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили _____

к.ф-м.н., доцент, Кустицкая Т.А

должность, инициалы, фамилия

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью изучения дисциплины является ознакомление студентов с основными моделями предиктивного и описательного статистического анализа.

1.2 Задачи изучения дисциплины

Задачами изучения дисциплины является:

- ознакомление с современными методами анализа статистических данных и сферами их применения в исследовательской практике;
- получение практических навыков применения изученных методов для анализа статистических данных с помощью языка программирования R.

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

| Код и наименование индикатора достижения компетенции | Запланированные результаты обучения по дисциплине |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ОПК-2: Способен обоснованно выбирать, дорабатывать и применять для решения исследовательских и проектных задач математические методы и модели, осуществлять проверку адекватности моделей, анализировать результаты, оценивать надежность и качество функционирования систем | |
| ОПК-2.1: Знать основные математические модели и методы решения исследовательских и проектных задач | математические основы моделей прогнозирования на основе данных, поиска закономерностей в данных |
| ОПК-2.2: Уметь осуществлять проверку адекватности моделей, анализировать результаты. | подбирать подходящие типы моделей для решения поставленных задач с учетом особенностей исследуемых данных навыками программной реализации базовых алгоритмов прогнозирования и поиска закономерностей в данных, навыками проверки адекватности построенных моделей |
| ОПК-2.3: Владеть методами оценки надежности и качества функционирования систем. | основные способы оценки достоверности результатов анализа вычислять метрики качества моделей, построенных по данным навыками программной реализации проверки гипотез и оценивания качества моделей |
| ОПК-3: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности | |

| | |
|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| ОПК-3.1: Знать принципы работы современных информационных технологий; | основные принципы работы с массивами данных в программных подуктах для анализа данных |
| ОПК-3.2: Уметь применять информационные технологии при решении профессиональных задач; | программировать алгоритмы анализа данных средствами языка программирования R |
| ОПК-3.3: Использовать аналитические и научные пакеты прикладных программ для решения профессиональных задач; | навыками программной предобработки данных, описательного анализа данных навыками визуализации результатов анализа данных |

1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется с применением ЭО и ДОТ

URL-адрес и название электронного обучающего курса: <https://e.sfu-kras.ru/course/view.php?id=22438>.

2. Объем дисциплины (модуля)

| Вид учебной работы | Всего, зачетных единиц (акад. час) | е |
|--------------------------------------------|---------------------------------------------|---|
| | | 1 |
| Контактная работа с преподавателем: | 2 (72) | |
| занятия лекционного типа | 1 (36) | |
| практические занятия | 1 (36) | |
| Самостоятельная работа обучающихся: | 1 (36) | |
| курсовое проектирование (КП) | Нет | |
| курсовая работа (КР) | Нет | |

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

| | | Контактная работа, ак. час. | | | | | | | |
|--------------------------------------------------------------|------------------------------------------------------------------------------|--------------------------------|--------------------------|-------------------------------------------|--------------------------|--------------------------------------------|--------------------------|-------------------------------------|--------------------------|
| № п/п | Модули, темы (разделы) дисциплины | Занятия лекционного типа | | Занятия семинарского типа | | | | Самостоятельная работа, ак. час. | |
| | | | | Семинары и/или Практические занятия | | Лабораторные работы и/или Практикумы | | | |
| | | Всего | В том числе в ЭИОС | Всего | В том числе в ЭИОС | Всего | В том числе в ЭИОС | Всего | В том числе в ЭИОС |
| 1. Описательная статистика и визуальный анализ данных | | | | | | | | | |
| | 1. Машинное обучение и анализ данных | 2 | | | | | | | |
| | 2. Основы программирования на языке статистических вычислений R | | | 2 | | | | | |
| | 3. Описательная статистика и визуальный анализ данных | 2 | | | | | | | |
| | 4. Описательная статистика и визуальный анализ данных | | | 2 | | | | | |
| | 5. Описательная статистика и визуальный анализ данных | | | | | | | 4 | |
| 2. Алгоритмы обучения с учителем | | | | | | | | | |
| | 1. Линейные модели обучения на размеченных данных: регрессия и классификация | 6 | | | | | | | |
| | 2. Линейные модели обучения на размеченных данных: регрессия и классификация | | | 6 | | | | | |

| | | | | | | | | |
|-------------------------------------------------|----|--|----|--|--|--|----|--|
| 3. Метрики качества | 2 | | | | | | | |
| 4. Метрики качества | | | 2 | | | | | |
| 5. Проблема переобучения | 2 | | | | | | | |
| 6. Проблема переобучения | | | 2 | | | | | |
| 7. Байесовские модели классификации и регрессии | 2 | | | | | | | |
| 8. Байесовские модели классификации и регрессии | | | 2 | | | | | |
| 9. Метрические алгоритмы | 2 | | | | | | | |
| 10. Метрические алгоритмы | | | 2 | | | | | |
| 11. Нейронные сети | 2 | | | | | | | |
| 12. Нейронные сети | | | 2 | | | | | |
| 13. Алгоритмы обучения с учителем | | | | | | | 16 | |
| 3. Алгоритмы обучения без учителя | | | | | | | | |
| 1. Задача кластеризации | 4 | | | | | | | |
| 2. Задача кластеризации | | | 4 | | | | | |
| 3. Метод главных компонент | 4 | | | | | | | |
| 4. Метод главных компонент | | | 4 | | | | | |
| 5. Выявление ассоциативных правил | 2 | | | | | | | |
| 6. Выявление ассоциативных правил | | | 2 | | | | | |
| 7. Алгоритмы обучения без учителя | | | | | | | 10 | |
| 4. Анализ связей | | | | | | | | |
| 1. Критерии различия | 3 | | | | | | | |
| 2. Критерии различия | | | 3 | | | | | |
| 3. Дисперсионный анализ | 3 | | | | | | | |
| 4. Дисперсионный анализ | | | 3 | | | | | |
| 5. Анализ связей | | | | | | | 6 | |
| Всего | 36 | | 36 | | | | 36 | |

4 Учебно-методическое обеспечение дисциплины

4.1 Печатные и электронные издания:

1. Ниворожкина Л. И., Арженовский С. В., Рудяга А. А., Торопова Н. А., Федосова О. Н., Житников И. В., Трегубова А. А., Федотова Э. А. Статистические методы анализа данных: Учебник(Москва: Издательский Центр РИО□).
2. Джеймс Г. Введение в статистическое обучение с примерами на языке R (Москва: ДМК Пресс).
3. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников(Москва: Физматлит).
4. Кабаков Р., Волкова П. А. R в действии. Анализ и визуализация данных на языке R(Москва: ДМК Пресс).
5. Волкова П. А., Шипунов А. Б. Статистическая обработка данных в учебно-исследовательских работах: Учебное пособие(Москва: Издательство "ФОРУМ").
6. Рубан А. И. Методы анализа данных: учебное пособие(Красноярск: СФУ).

4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):

1. Для выполнения индивидуальных заданий по программированию алгоритмов анализа данных необходима программная среда вычислений R версии не младше, чем 3.4.3, а также редактор кода RStudio версии 1.1 или старше.

4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:

1. Научная электронная библиотека открытого доступа «Киберленинка»
URL: <http://cyberleninka.ru/>

5 Фонд оценочных средств

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

Для проведения лекционных занятий необходим мультимедийный проектор
Практические занятия проводятся в компьютерном классе