

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

Б1.В.ДВ.04.01 Анализ неструктурированных данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

27.03.03 СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ

Направленность (профиль)

27.03.03 СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ

Форма обучения

очная

Год набора

2019

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

Программу составили \_\_\_\_\_

\_\_\_\_\_ канд. техн. наук, доцент, Даничев Алексей Александрович

\_\_\_\_\_ должность, инициалы, фамилия

## 1 Цели и задачи изучения дисциплины

### 1.1 Цель преподавания дисциплины

Целью преподавания дисциплины "Аналитическая обработка структурированных и неструктурированных данных является изучение основных задач и методов обработки и анализа данных и освоение методов статистического анализа при работе с неструктурированными данными.

### 1.2 Задачи изучения дисциплины

Для достижения цели изучения дисциплины предполагается решение студентами следующих задач:

- изучение методов статистического анализа данных;
- изучение методов систематизации информации неструктурированных данных;
- изучение принципов организации и проведения аналитического исследования;
- решение практических задач, прогнозирование и выработка рекомендаций.

### 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  |
|---|--|
| <b>ОПК-2: способностью применять аналитические, вычислительные и системно-аналитические методы для решения прикладных задач в области управления объектами техники, технологии, организационными системами, работать с традиционными носителями информации, базами знаний</b> |  |
| ОПК-2: способностью применять аналитические, вычислительные и системно-аналитические методы для решения прикладных задач в области управления объектами техники, технологии, организационными системами, работать с традиционными носителями информации, базами знаний        | Знать теоретические основы моделирования, методологию построения моделей сложных систем, методы сбора, хранения и анализа данных.<br>Уметь применять системы анализа неструктурированной информации для решения задач сквозного поиска по источникам, выявления закономерностей на основании анализа текстовых данных, извлечения ключевых факторов из неструктурированных текстов.<br>Владеть методами теории систем и системного анализа |
| <b>ОПК-7: способностью к освоению новой техники, новых методов и новых технологий</b>   |  |

|   |   |
|---|---|
| ОПК-7: способностью к освоению новой техники, новых методов и новых технологий  | Знать методы сбора данных<br>Уметь применять способности по выявлению и прогнозированию основных направлений использования анализа неструктурированных данных<br>Владеть научными исследованиями и готовить аналитические материалы для оценки мероприятий и  |
|   | выработки стратегических решений  |
| <b>ПК-1: способностью принимать научно-обоснованные решения на основе математики, физики, химии, информатики, экологии, методов системного анализа и теории управления, теории знаний, осуществлять постановку и выполнять эксперименты по проверке их корректности и эффективности</b> |   |
| ПК-1: способностью принимать научно-обоснованные решения на основе математики, физики, химии, информатики, экологии, методов системного анализа и теории управления, теории знаний, осуществлять постановку и выполнять эксперименты по проверке их корректности и эффективности        | Знать характеристики рынка систем анализа неструктурированных данных и перспективы развития сегмента информационно-технологической отрасли, получившей название «Большие данные» (Big Data), основные методы анализа, применяемые в «Больших данных».<br>Уметь применять способности для обоснования решения на основе результатов моделирования<br>Владеть методами систематизации информации неструктурированных 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=22468>.

## 2. Объем дисциплины (модуля)

| Вид учебной работы                         | Всего,<br>зачетных<br>единиц<br>(акад.час) | е |
|--|--|---|
|  |  | 1 |
| <b>Контактная работа с преподавателем:</b> | <b>3 (108)</b>                             |   |
| занятия лекционного типа                   | 1 (36)                                     |   |
| практические занятия                       | 2 (72)                                     |   |
| <b>Самостоятельная работа обучающихся:</b> | <b>3 (108)</b>                             |   |
| курсовое проектирование (КП)               | Нет  |   |
| курсовая работа (КР)                       | Нет  |   |
| <b>Промежуточная аттестация (Экзамен)</b>  | <b>1 (36)</b>                              |   |

### 3 Содержание дисциплины (модуля)

#### 3.1 Разделы дисциплины и виды занятий (тематический план занятий)

| №<br>п/п                              |  | Модули, темы (разделы) дисциплины             |  | Контактная работа, ак. час.    |  |   |                          |  |                          |                                     |  |
|---------------------------------------|--|---|--|--------------------------------|--|---|--------------------------|--|--------------------------|-------------------------------------|--|
|                                       |  |   |  | Занятия<br>лекционного<br>типа |  | Занятия семинарского типа                 |                          |  |                          | Самостоятельная<br>работа, ак. час. |  |
|                                       |  |   |  |                                |  | Семинары и/или<br>Практические<br>занятия |                          | Лабораторные<br>работы и/или<br>Практикумы |                          |                                     |  |
|                                       |  |   |  |                                |  | Всего                                     | В том<br>числе в<br>ЭИОС | Всего                                      | В том<br>числе в<br>ЭИОС |                                     |  |
| <b>1. Анализ текста</b>               |  |   |  |                                |  |   |                          |  |                          |                                     |  |
|                                       |  | 1. Способы анализа текстовой информации       |  | 2                              |  |   |                          |  |                          |                                     |  |
|                                       |  | 2. Стеммер, лематизатор                       |  |                                |  | 4   |                          |  |                          |                                     |  |
|                                       |  | 3. Обработка текста для машинного обучения    |  | 2                              |  |   |                          |  |                          |                                     |  |
|                                       |  | 4. Формирование n-грамм                       |  |                                |  | 4   |                          |  |                          |                                     |  |
|                                       |  | 5. Определение тональности текста             |  | 2                              |  |   |                          |  |                          |                                     |  |
|                                       |  | 6. Определение тональности текста             |  |                                |  | 4   |                          |  |                          |                                     |  |
|                                       |  | 7. Способы анализа текстовой информации       |  |                                |  |   |                          |  |                          | 36                                  |  |
| <b>2. Структурирование информации</b> |  |   |  |                                |  |   |                          |  |                          |                                     |  |
|                                       |  | 1. Кластеризация и классификация              |  | 2                              |  |   |                          |  |                          |                                     |  |
|                                       |  | 2. Определение тематики текста                |  |                                |  | 4   |                          |  |                          |                                     |  |
|                                       |  | 3. Численная оценка качества работы алгоритма |  | 2                              |  |   |                          |  |                          |                                     |  |
|                                       |  | 4. Численная оценка качества работы алгоритма |  |                                |  | 4   |                          |  |                          |                                     |  |
|                                       |  | 5. Ансамбль методов                           |  | 2                              |  |   |                          |  |                          |                                     |  |

|  |   |  |   |  |  |  |    |  |
|--|---|--|---|--|--|--|----|--|
| 6. Ансамбль методов  |   |  | 4 |  |  |  |    |  |
| 7. Факторизация  | 2 |  |   |  |  |  |    |  |
| 8. Факторизация  |   |  | 4 |  |  |  |    |  |
| 9. Нейронные сети  | 2 |  |   |  |  |  |    |  |
| 10. Нейронные сети   |   |  | 4 |  |  |  |    |  |
| 11. Нечеткая логика  | 2 |  |   |  |  |  |    |  |
| 12. Нечеткая логика  |   |  | 4 |  |  |  |    |  |
| 13. Структурирование информации                                    |   |  |   |  |  |  | 36 |  |
| <b>3. Элементы теории вероятностей и математической статистики</b> |   |  |   |  |  |  |    |  |
| 1. Библиотеки для статистических расчётов                          | 2 |  |   |  |  |  |    |  |
| 2. Нормализация данных   |   |  | 4 |  |  |  |    |  |
| 3. Случайные величины. Характеристики выборок                      | 2 |  |   |  |  |  |    |  |
| 4. Построение валидных гистограмм                                  |   |  | 4 |  |  |  |    |  |
| 5. Устойчивые распределения  | 2 |  |   |  |  |  |    |  |
| 6. Закон Бенфорда  |   |  | 4 |  |  |  |    |  |
| 7. Проверка стат. гипотез  | 2 |  |   |  |  |  |    |  |
| 8. Проверка гипотез  |   |  | 4 |  |  |  |    |  |
| 9. Непараметрические критерии                                      | 2 |  |   |  |  |  |    |  |
| 10. Непараметрические критерии                                     |   |  | 4 |  |  |  |    |  |
| 11. Проверка закона распределения СВ                               | 2 |  |   |  |  |  |    |  |
| 12. Проверка закона распределения СВ                               |   |  | 4 |  |  |  |    |  |
| 13. Связные выборки  | 2 |  |   |  |  |  |    |  |
| 14. Связные выборки  |   |  | 4 |  |  |  |    |  |
| 15. Анализ мощности  | 2 |  |   |  |  |  |    |  |
| 16. Предсказание ошибки эксперимента                               |   |  | 4 |  |  |  |    |  |

|   |    |  |    |  |  |  |     |  |
|---|----|--|----|--|--|--|-----|--|
| 17. Планирование эксперимента                   | 2  |  |    |  |  |  |     |  |
| 18. Интерпретация результатов эксперимента      |    |  | 4  |  |  |  |     |  |
| 19. Планирование эксперимента, проверка гипотез |    |  |    |  |  |  | 36  |  |
| Всего   | 36 |  | 72 |  |  |  | 108 |  |



#### **4 Учебно-методическое обеспечение дисциплины**

##### **4.1 Печатные и электронные издания:**

1. Хайкин С. Нейронные сети: полный курс(Санкт-Петербург: Вильямс).
2. Шевелева И. В., Васильева А. В. Теория вероятностей и математическая статистика: методические указания по выполнению контрол. работ для студентов заоч. формы обучения(Красноярск: ИПК СФУ).
3. Сидняев Н. И., Вилисова Н. Т. Введение в теорию планирования эксперимента: учебное пособие для вузов по направлению "Машиностроительные технологии и оборудование" специальности "Реновация средств и объектов материального производства в машиностроении"(Москва: МГТУ имени Н.Э.Баумана).
4. Даничев А. А. Компьютерные технологии поддержки принятия решений: учеб.-метод. пособие для лаб. работ студентов направления 220100.62 «Системный анализ и управление».(Красноярск: СФУ).
5. Помазан В. А., Ковалевич И. А., Шестаков В. Н., Даничев А. А. Информационные технологии в управлении человеческими ресурсами: электрон. учеб.-метод. комплекс дисциплины(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. Любой инструментарий для статистического анализа.

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

1. Не требуется

#### **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)**

Учебная аудитория для проведения лекционных и практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещение для самостоятельной работы, содержащее специализированную мебель, компьютеры с подключением к сети Интернет (неограниченный доступ) и обеспечением доступа в электронную информационно-образовательную среду Университета, доступом к системе виртуальных машин; демонстрационное оборудование: интерактивная доска обратной проекции, проектор, экран для проектора; доступ к беспроводной сети WI-FI, маркерная доска.