

**Министерство образования и науки Российской Федерации**  
федеральное государственное автономное образовательное учреждение  
высшего образования  
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики  
Направление подготовки 09.04.03. Прикладная информатика  
Кафедра Программная инженерия

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

| Тема работы                                                           |
|-----------------------------------------------------------------------|
| Система поддержки медицинских научных исследований бронхиальной астмы |

УДК 616.248:001.891

Студент

| Группа | ФИО                        | Подпись | Дата |
|--------|----------------------------|---------|------|
| 8KM51  | Сеидова Айсель Султан кызы |         |      |

Руководитель

| Должность      | ФИО           | Ученая степень,<br>звание | Подпись | Дата |
|----------------|---------------|---------------------------|---------|------|
| Доцент каф. ПИ | Марухина О.В. | К. Т. Н.                  |         |      |

**КОНСУЛЬТАНТЫ:**

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

| Должность          | ФИО           | Ученая степень,<br>звание | Подпись | Дата |
|--------------------|---------------|---------------------------|---------|------|
| Ассистент каф. МЕН | Баннова К. А. | К. Э. Н.                  |         |      |

По разделу «Социальная ответственность»

| Должность       | ФИО                | Ученая степень,<br>звание | Подпись | Дата |
|-----------------|--------------------|---------------------------|---------|------|
| Доцент каф. ЭБЖ | Пустовойтова М. И. | К. Х. Н.                  |         |      |

**ДОПУСТИТЬ К ЗАЩИТЕ:**

| Зав. кафедрой | ФИО          | Ученая степень,<br>звание | Подпись | Дата |
|---------------|--------------|---------------------------|---------|------|
| Зав. каф. ПИ  | Иванов М. А. | К. Т. Н.                  |         |      |

## Запланированные результаты обучения по программе

| Код<br>результатов | Результат обучения<br>(выпускник должен быть готов)                                                                                                                                                                                                                                                                                                   |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                    | <b>Профессиональные компетенции</b>                                                                                                                                                                                                                                                                                                                   |
| Р1                 | Применять базовые и специальные естественно-научные и математические знания в области информатики, экономики, маркетинга и менеджмента, достаточные для комплексной инженерной деятельности.                                                                                                                                                          |
| Р2                 | Применять базовые и специальные знания в области современных информационных технологий для решения инженерных и экономических задач.                                                                                                                                                                                                                  |
| Р3                 | Ставить и решать задачи комплексного анализа, связанные с созданием новых информационных технологий и информационных систем в экономике, с использованием базовых и специальных знаний, современных аналитических методов и моделей.                                                                                                                  |
| Р4                 | Разрабатывать новые и модернизировать уже существующие информационные технологии и системы (в экономике) в соответствии с техническим заданием.                                                                                                                                                                                                       |
| Р5                 | Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области прикладной информатики. Проводить исследования, связанные с оценкой информационной безопасности проектов. |
| Р6                 | Внедрять, эксплуатировать и обслуживать современные информационные технологии и системы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.                                                                                                               |
|                    | <b>Универсальные компетенции</b>                                                                                                                                                                                                                                                                                                                      |
| Р7                 | Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.                                                                                                                                                                                                                           |
| Р8                 | Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.                                                                                                                                                                    |
| Р9                 | Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.                                                                                                          |
| Р10                | Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.                                                                                                                                                                                                                                 |
| Р11                | Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.                                                                                                                                                                                                             |

**Министерство образования и науки Российской Федерации**  
федеральное государственное автономное образовательное учреждение  
высшего образования  
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Кибернетики

Направление подготовки (специальность) 09.04.03. Прикладная информатика

Кафедра Программная инженерия

УТВЕРЖДАЮ:

Зав. кафедрой ПИ

\_\_\_\_\_  
(Подпись)      \_\_\_\_\_  
(Дата)      Иванов М.А  
(Ф.И.О.)

**ЗАДАНИЕ**

**на выполнение выпускной квалификационной работы**

В форме:

|                                                                            |
|----------------------------------------------------------------------------|
| магистерской диссертации                                                   |
| (бакалаврской работы, дипломного проекта/работы, магистерской диссертации) |

Студенту:

| Группа | ФИО                         |
|--------|-----------------------------|
| 8KM51  | Сеидовой Айсель Султан кызы |

Тема работы:

|                                                                       |                      |
|-----------------------------------------------------------------------|----------------------|
| Система поддержки медицинских научных исследований бронхиальной астмы |                      |
| Утверждена приказом директора (дата, номер)                           | 1484/с от 01.03.2017 |

Срок сдачи студентом выполненной работы:

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| <b>Исходные данные к работе</b><br>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).        | Клинико-лабораторные показатели городской клинической больницы №3                                           |
| <b>Перечень подлежащих исследованию, проектированию и разработке вопросов</b><br>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе). | Постановка задачи<br>Выбор решения задачи<br>Решение поставленной задачи<br>Описание полученных результатов |
| <b>Перечень графического материала</b><br>(с точным указанием обязательных чертежей)                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                                             |
| <b>Консультанты по разделам выпускной квалификационной работы</b><br>(с указанием разделов)                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                             |

| Раздел                                                                                  | Консультант                  |
|-----------------------------------------------------------------------------------------|------------------------------|
| Разделы 1-3                                                                             | Марухина Ольга Владимировна  |
| Раздел 4                                                                                | Баннова Кристина Алексеевна  |
| Раздел 5                                                                                | Пустовойтова Марина Игоревна |
| <b>Названия разделов, которые должны быть написаны на русском и иностранном языках:</b> |                              |
| Введение                                                                                |                              |
| Глава 1 Постановка задачи                                                               |                              |
| Глава 2 Выбор решения задачи                                                            |                              |

|                                                                                                 |  |
|-------------------------------------------------------------------------------------------------|--|
| <b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b> |  |
|-------------------------------------------------------------------------------------------------|--|

**Задание выдал руководитель:**

| Должность      | ФИО            | Ученая степень, звание | Подпись | Дата |
|----------------|----------------|------------------------|---------|------|
| Доцент каф. ПИ | Марухина О. В. | к. т. н.               |         |      |

**Задание принял к исполнению студент:**

| Группа | ФИО                        | Подпись | Дата |
|--------|----------------------------|---------|------|
| 8KM51  | Сеидова Айсель Султан кызы |         |      |

## **РЕФЕРАТ**

Выпускная квалификационная работа 91 с., 11 рис., 22 табл., 28 источников, 2 прил.

Ключевые слова: многомерные данные, Data Mining, закономерности, кластеризация, логические правила.

Объект исследования – многомерные данные (в частности, клинико-лабораторные показатели пациентов с бронхиальной астмой).

Предмет исследования – обработка многомерных данных с использованием пакетов языка R.

Целью работы является создание системы поддержки медицинских научных исследований бронхиальной астмы, на основе анализа клинико-лабораторных показателей.

Метод исследования – теоретический, практический, получение информации о предметной области, её исследование и анализ.

В результате анализа предметной области, проведен анализ клинико-лабораторных показателей, на основе полученных результатов сформирована база знаний.

Выпускная квалификационная работа выполнена в текстовом редакторе Microsoft Word 2010 и представлена на диске (в конверте на обороте обложки).

## Оглавление

|                                                                                                                                                |    |
|------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Введение.....                                                                                                                                  | 8  |
| Обзор литературы.....                                                                                                                          | 10 |
| Глава 1 Постановка задачи .....                                                                                                                | 12 |
| 1.1. Понятие информационной системы.....                                                                                                       | 12 |
| 1.2 Описание предметной области .....                                                                                                          | 12 |
| 1.4 Формальная постановка задачи .....                                                                                                         | 15 |
| Глава 2 Выбор методов решения задачи .....                                                                                                     | 16 |
| 2.1 Выбор технологии и описание методов решения задачи .....                                                                                   | 16 |
| 2.1.1 База знаний .....                                                                                                                        | 16 |
| 2.1.2 Стратегии получения знаний .....                                                                                                         | 16 |
| 2.1.3 Технология Data Mining.....                                                                                                              | 17 |
| 2.1.4. Кластерный анализ.....                                                                                                                  | 17 |
| 2.1.5. Построение деревьев решений.....                                                                                                        | 18 |
| 2.2. Выбор инструмента решения.....                                                                                                            | 19 |
| Глава 3 Решение поставленной задачи и описание полученных результатов .....                                                                    | 23 |
| 3.1. Подготовка данных к анализу .....                                                                                                         | 23 |
| 3.2. Алгоритм выявления скрытых закономерностей .....                                                                                          | 23 |
| 3.3. Кластерный анализ данных.....                                                                                                             | 24 |
| 3.4. Построение деревьев решений .....                                                                                                         | 29 |
| 3.5. Анализ полученных результатов.....                                                                                                        | 34 |
| Глава 4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....                                                                   | 35 |
| 4.1. Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения..... | 35 |
| 4.1.1. Потенциальные потребители результатов исследования.....                                                                                 | 35 |
| 4.1.2. Анализ конкурентных технических решений .....                                                                                           | 35 |
| 4.1.3. Технология QuaD .....                                                                                                                   | 37 |
| 4.1.4. SWOT-анализ.....                                                                                                                        | 38 |
| 4.2. Планирование научно-исследовательских работ .....                                                                                         | 39 |
| 4.2.1. Структура работ в рамках научного исследования .....                                                                                    | 39 |
| 4.2.2. Определение трудоемкости выполнения работ .....                                                                                         | 40 |
| 4.2.3. Разработка графика проведения научного исследования .....                                                                               | 42 |
| 4.3. Бюджет научно-технического исследования (НТИ) .....                                                                                       | 47 |
| 4.3.1. Расчет материальных затрат НТИ .....                                                                                                    | 47 |
| 4.3.2. Основная заработная плата исполнителей темы .....                                                                                       | 47 |

|                                                                                                                                    |    |
|------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.3.3. Дополнительная заработная плата.....                                                                                        | 49 |
| 4.3.4. Отчисления во внебюджетные фонды (страховые отчисления) .....                                                               | 49 |
| 4.3.5. Накладные расходы.....                                                                                                      | 50 |
| 4.3.6. Формирование бюджета затрат научно-исследовательского проекта.....                                                          | 50 |
| 4.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования..... | 51 |
| Глава 5 Социальная ответственность .....                                                                                           | 56 |
| 5.1. Техногенная безопасность .....                                                                                                | 56 |
| 5.2. Региональная безопасность.....                                                                                                | 62 |
| 5.3. Организационные мероприятия обеспечения безопасности .....                                                                    | 63 |
| 5.4. Особенности законодательного регулирования проектных решений.....                                                             | 63 |
| 5.5. Безопасность в чрезвычайных ситуациях. ....                                                                                   | 64 |
| Заключение.....                                                                                                                    | 66 |
| Список публикаций студента .....                                                                                                   | 68 |
| Список использованных источников.....                                                                                              | 72 |
| Приложение А.....                                                                                                                  | 75 |
| Приложение Б .....                                                                                                                 | 78 |

## **Введение**

Бронхиальная астма (БА) является глобальной проблемой, с каждым годом ее актуальность во всем мире возрастает. БА также называют одной из болезней цивилизации, так как ее распространенность в современном обществе постоянно растет. Статистика показывает, что у 60-80% взрослых пациентов данное заболевание появилось в детском возрасте.

Резкий рост заболеваемости астмой произошел в конце XX века. В 1998 году число больных бронхиальной астмой в мире достигало 155 млн человек, а в настоящее время эта цифра превышает 300 млн. Исходя из данных некоторых прогностических аналитических исследований, можно сделать вывод, что к 2025 году, в том случае, если процесс урбанизации будет продолжаться теми же темпами, бронхиальная астма разовьется дополнительно у 100–150 млн человек [1]. Согласно данным статистических материалов МЗ РФ за 2014 г., в России официально зарегистрировано 1406493 больных бронхиальной астмой. По оценкам специалистов, численность больных астмой как минимум в 5–6 раз превышает данные официальной статистики и составляет приблизительно 9,915 млн [2].

Накопленные в архивах медицинские данные содержат большой запас информации о различных случаях заболеваний, о тяжести и ходе лечения, а также хранят огромное количество клинико-лабораторных показателей.

По данным Федерального информационного фонда данных социально-гигиенического мониторинга (2014 г.) Томская область отнесена к территориям «риска» по уровню заболеваемости населения астмой. В 2014 г. заболеваемость астмой (астматический статус) населения Томской области в возрастных группах находилась на следующих уровнях: дети – 238,2 на 100 тыс. населения данной возрастной группы, подростки – 142,9, взрослые – 107,8.

В связи с этим актуальность темы выпускной квалификационной работы не вызывает сомнений.

Решением проблемы БА может послужить анализ клинико-лабораторных показателей, которые отражают состояние и клиническую картину и нахождение закономерностей, помогающих корректировать лечение.

Целью работы является создание системы поддержки медицинских научных исследований бронхиальной астмы, на основе анализа клинико-лабораторных показателей.

Выпускная квалификационная работа посвящена решению следующих задач:

- выявить скрытые закономерности эффективности лечения пациентов;

- сформировать группы пациентов со сходными значениями клинико-лабораторных показателей;

- на основании выявленных скрытых закономерностей, сформировать базу знаний;

#### **Внедрение результатов.**

1. Участие в выполнении работ по всероссийскому гранту: РФФИ 14-07-00675 «Система поддержки медицинских научных исследований психогенных форм бронхиальной астмы» (под рук. профессора каф. ПИ О.Г. Берестневой).

2. Публикация в журнале «Современные наукоемкие технологии»: Жаркова О. С. , Шаропин К. А. , Сеидова А. С. , Берестнева Е. В. , Осадчая И. А. Построение систем поддержки принятия решений в медицине на основе деревьев решений // Современные наукоемкие технологии. - 2016 - №. 6-1. - С. 33-37

3. Участие в Международной научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР–2017» с докладом: «Инструментальные методы анализа медицинских данных».

4. Участие в XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии» с докладом: «Построение СППР на основе деревьев решений в медицинских исследованиях».

5. Участие в Международной научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР–2016» с докладом: «Формирование базы знаний для медицинской информационной системы».

6. Участие в XXII Международной научно-технической конференции, Нижний Новгород «Информационные системы и технологии (ИСТ–2016)» с докладом: «Использование пакета WIZWHY для формирования базы знаний медицинской экспертной системы»

7. Участие в III Международной конференции «Информационные технологии в науке, управлении, социальной сфере и медицине» с докладом: «Построение СППР в медицине на основе деревьев решений».

8. Участие в III Международной конференции «Информационные технологии в науке, управлении, социальной сфере и медицине» с докладом: «Использование пакета WIZWHY для формирования базы знаний экспертных систем».

9. Участие в XIII Международной научно- практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии» с докладом «Формирование базы знаний на основе пакета WIZWHY для медицинской информационной системы».

## Обзор литературы

Информационные системы активно используются в медицине, одним из важных направлений является создание на их основе систем поддержки принятия решений (СППР). СППР используют для диагностики и прогнозирования заболеваний. Подобные системы не могут нести ответственность за принятые с ее помощью решения, но СППР способны существенно упростить и ускорить работу врача [3].

Таким образом, важными условиями создания СППР являются понятность и простота методов ее построения для врача. Использование различных статистических критериев, подтверждающих либо опровергающих выдвинутые врачом гипотезы позволяет добиться необходимого результата.

Среди информационных и интеллектуальных систем, связанных с проблемами диагностики и изучения бронхиальной астмы можно выделить следующие:

1. СППР для оценки контролируемости бронхиальной астмы по результатам суточной пикфлоуметрии (Дальневосточный научный центр физиологии и патологии дыхания СО РАМН (Н.С. Безруков, В.П. Колосов, Ю.М. Перельман, Ю.Ю. Хижняк)). Система состоит из двух подсистем, работающих независимо друг от друга, признаки которых выбираются медиком-экспертом при помощи статистических критериев. Реализована на базе пакета «Medical Toolbox».

2. Система поддержки принятия решения для диагностики бронхиальной астмы по нейрофизиологическим параметрам на основе адаптивной нейро-нечеткой сети /Дальневосточный научный центр физиологии и патологии дыхания СО РАМН (Н.С. Безруков, Е.В. Ермакова, В.П. Колосов, Ю.М. Перельман). При диагностике используются параметры, значимость которых подтверждена статистическими критериями. На основе значимых параметров строятся подсистемы диагностики в пакете «Medical Toolbox».

3. Web-сайт «Виртуальный Астма-центр» и его версия для мобильных устройств (Хузина Е.А., Фурман Е.Г., Малинин С.В., Корюкина И.П.)

4. Проблемно-ориентированная информационная система лечебно-профилактической помощи больных бронхиальной астмой (Воронежская государственная медицинская академия им. Н.Н. Бурденко (А.В. Будневский, В.Т. Бурлачук, А.В. Разворотнев)).

Следует отметить, что медицинские базы знаний могут получить доверие врачей, только в случае, если они будут развиваться коллективно медицинским сообществом. Развитие глобальной мировой компьютерной сети Интернет создало возможности для коллективного развития и использования информационных ресурсов, в том числе баз знаний и экспертных систем. В качестве примера можно привести многоцелевой банк

знаний (Институт автоматизации и процессов управления ДВО РАН), предназначенный для создания и развития совокупности специализированных банков знаний, одним из которых является банк знаний в области медицинской диагностики. Многоцелевой банк знаний предназначен для поддержки медицинского образования, медицинских научных исследований и дистанционных консультаций по диагностике заболеваний.

На сегодняшний день в Институте автоматизации и процессов управления ДВО РАН используется предложенное Д.В.Нагорным и М.Ю. Черняховской формальное представление базы наблюдений для компьютерного банка медицинских знаний. Данный подход был использован при реализации проекта, поскольку представляется нам наиболее эффективным.

В проектируемой в рамках ВКР системе поддержки медицинских исследований форм бронхиальной астмы использован опыт создания аналогичных МИС (в частности, перечисленных выше).

Несмотря на большое количество исследований, посвященных проблеме роли психосоциальных факторов в генезе бронхиальной астмы, их результаты непоследовательные и противоречивые (Douwes J., Brooks C., and Pearce N. 2011). А связи между психосоциальными факторами и астмой и другими атопическими расстройствами носят двусторонний характер, при котором не всегда можно различить вклад поведенческих или физиологических путей (Chida Y, Hamer M, Steptoe A. 2008).

## **Глава 1 Постановка задачи**

### **1.1. Понятие информационной системы**

Информационная система (ИС) - это взаимосвязанная совокупность средств, методов и персонала, используемых для хранения, обработки и выдачи информации в интересах достижения поставленной цели [4].

Данный термин используется как в широком, так и в узком смысле.

В широком смысле информационная система - это совокупность технического, программного и организационного обеспечения, а также персонала, которая предназначена для своевременного обеспечения надлежащих людей необходимой информацией.

В узком смысле информационная система - это подмножество компонентов ИС в широком смысле, которое включает в себя базы данных, СУБД и специализированные прикладные программы. ИС в узком смысле рассматривается как программно-аппаратная система, которая предназначена для автоматизации целенаправленной деятельности конечных пользователей, обеспечивающую, в соответствии с заложенной в неё логикой обработки, возможность получения, модификации и хранения информации.

### **1.2 Описание предметной области**

Городской клинической больницей №3 предоставлен массив данных, в котором содержатся группы клинико-лабораторных показателей пациентов: до лечения и после лечения. Пациенты – взрослые, страдающие БА. Решением задачи является формирование базы знаний на основе имеющихся показателей. База знаний в дальнейшем будет использоваться в системе поддержки медицинских исследований форм бронхиальной астмы.

Астма классифицируется по иммунному механизму (атопическая и неатопическая), по типу воспаления (эозинофильная и нейтрофильная), по особенностям клиники (ночная астма, астма физического усилия) по ответу на терапию (хорошо контролируемая, неконтролируемая, и т.д. И именно так называемая «трудная астма», резистентная к стероидной терапии, вновь возродила интерес исследователей к типированию этой болезни (.Haldar P, Pavord ID, Shaw DE, et al. 2008; Wenzel SE, Schwartz LB, Langmack EL, et al.1999.) Возникла необходимость идентифицировать фенотипы с определенными патофизиологическими отклонениями, чтобы фокусировать (сосредотачивать) текущие генетические и молекулярные исследования на разработку новых видов молекулярной терапии, которые будут эффективны в специфических подгруппах астмы (Heaney LG, Robinson DS. 2005.) Противоречивость результатов исследований психологического направления возможно связано с тем, что все больные БА расцениваются ими как

однородная популяция людей в плане соматического статуса, но с разными психологическими состояниями, не учитывая их клинических особенностей. С другой стороны, клиницисты (пульмонологи, терапевты) не всегда придают значение факту, что разные эмоциональные состояния и психические расстройства влекут за собой различные физиологические ответы у здоровых и больных астмой. То есть многообразие психологических воздействий вызывает многообразие психологических и соматических изменений в различных группах больных БА. Поэтому необходимо изучение психологических (психических) и социальных факторов в тесной взаимосвязи с клиническими проявлениями и особенностями манифестации и течения болезни на основе современной парадигмы БА (хроническое воспаление и гиперреактивность дыхательных путей) с междисциплинарных позиций.

В основе исследования лежит предложенная томскими медиками (Языков К.Г., Немеров Е.В.) психосинергетическая модель психогенной бронхиальной астмы [5]. В соответствии с данной моделью выделены эмпирические группы условно названные:

БАСП - бронхиальная астма соматопсихогенная (предварительный термин);

БАПИ - бронхиальная астма психогенно - индуцированная;

БАНП - бронхиальная астма непсихогенная;

ПО – психогенная одышка.

В группу, условно названную БА психогенно-индуцированная (БАПИ), вошли пациенты, у которых ведущий элемент болезни, первый приступ удушья развился после перенесенного эмоционального стресса, психотравмирующего жизненного события. Дальнейшее резкое ухудшение течения болезни было связано с психологическими проблемами негативного характера. Вторую группу пациентов с БА непсихогенной (БАНП) составили лица с «классической» БА, преимущественно атопической формой заболевания, у которых в начале болезни наблюдались различные проявления атопии (риниты, конъюнктивиты, кожные высыпания). Обострение болезни чаще всего вызывали аллергии, вирусные инфекции, а не психологические факторы. Третья группа, БАСП, характеризуется нарушением «обычного» течения болезни в связи с жизненным стрессом, после чего психо - эмоциональные триггеры вызывали тяжелые приступы удушья, обострения болезни. Психогенная одышка – наиболее характерное и самое частое проявление невротических нарушений дыхания, встречающееся при панических расстройствах, астено-депрессивных и истерических состояниях, соматоформных дисфункциях вегетативной нервной системы и различных психосоматических заболеваниях.

В данной работе приняли участие 83 пациента.

Задача решается на основе исходных данных содержащих следующие показатели:

Код группы – обозначение (цифрами или буквами) той или иной группы пациентов;

Пол – пол больного

Возраст – возраст больного, лет

Рост – рост больного, см

Вес – вес больного, кг

Степ.ВН — показатель степень вентиляционной недостаточности

Возр.1эп. б-ни — возраст больного во время первого эпизода болезни, лет

Ст.тяж.БА — степень тяжести БА

Прод.б-ни — продолжительность болезни (лет)

Кл.пр.атопии – клинические проявления (признаки) атопии

**Показатели вентиляции легких:** 1-до, 2- после АВС

МОД – Минутный объем дыхания, л/мин

ЖЕЛ – Жизненная Емкость Легких, % (к **должным величинам**)

ФЖЕЛ – Форсированная ЖЕЛ, %

ОФВ1 – Объем Форсированного Выдоха за 1 секунду, %

МВЛ – Максимальная Вентиляция Легких за 1 мин, %

ОФВ1/ЖЕЛ – отношение ОФВ1/ЖЕЛ, %

ПОС выд. – Пиковая Объемная Скорость выдоха, %

МОС25 – Максимальная Объемная Скорость выдоха на уровне 25% ФЖЕЛ

МОС50 – Максимальная Объемная Скорость выдоха на уровне 50% ФЖЕЛ

МОС75 – Максимальной Объемная Скорость выдоха на уровне 75% ФЖЕЛ

Все показатели вентиляции легких (кроме МОД) выражены в % к **должным величинам**

**Показатели механики дыхания:** 1-до, 2- после АВС

C<sub>dyn</sub> – динамическая растяжимость (C) легких, л/см.вод.ст.

C<sub>stat</sub> – статическая растяжимость C) легких, л/см.вод.ст.

R<sub>вд</sub> — бронхиальное сопротивления на вдохе, см.вод.ст./л/сек

R<sub>выд</sub> – бронхиальное сопротивление на выдохе, см.вод.ст./л/сек

W<sub>общ</sub> – общая работа дыхания при спонтанном дыхании, кгм/мин

W<sub>уд</sub> – удельной работы дыхания (на литр вентиляции), кгм/л

W<sub>н.эл</sub> – неэластическая фракция общей работы дыхания, кгм/мин

W<sub>эл</sub> – эластическая фракция общей работы дыхания, кгм/мин

W<sub>МОД10</sub> – общая работа дыхания при МОД10, кгм/мин

W<sub>МОД15</sub> – общая работа дыхания при МОД15, кгм/мин

$W_{\text{МВЛобщ}}$  –общая работа дыхания при МВЛ, кгм/мин

$W_{\text{МВЛуд}}$  –работы дыхания при МВЛ, кгм/л

Показатели механики дыхания в дальнейшем приводятся к системе СИ

В ходе выполнения работы показатели рассматриваются как до лечения, так и после проведенного лечения, а также разность показателей. Необходимо создать систему поддержки медицинских исследований форм БА, которая бы позволила обосновать или наоборот опровергнуть существование 4 групп БА, которые были выявлены томскими медиками.

#### 1.4 Формальная постановка задачи

Исходный массив данных нашей задачи – матрица  $X_{nm}$ ,

где  $n$  – количество пациентов ( $n=83$ );

$m$  – число медицинских показателей ( $m=48$ ).

В решении задачи применяется многомерный анализ данных. Предметом многомерного анализа являются сложные системы, элементы которых характеризуются множеством зависимых между собой объектов и признаков, представляемые в виде матрицы, строки которой соответствуют наблюдаемым объектам, а столбцы – характеризующим их признаком.

$$\dot{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}, \quad (1)$$

где  $n$  – число объектов (строк);

$m$  – число признаков (столбцов);

$x_{nm}$  – конкретное значение признака  $m$  у  $k$ -го элемента  $x$   $n$ -го объекта.

Для решения поставленной задачи необходимо проанализировать существующие подходы к анализу многомерных данных, выбрать метод и способ его компьютерной реализации, который позволяет обрабатывать большие массивы данных и выявлять скрытые знания.

В результате решения задачи будет построена система поддержки медицинских исследований БА, основанная на методах интеллектуального анализа данных.

## **Глава 2 Выбор методов решения задачи**

### **2.1 Выбор технологии и описание методов решения задачи**

#### **2.1.1 База знаний**

База знаний – это особого рода база данных, разработанная для управления знаниями, то есть сбором, хранением, поиском и выдачей знаний [5]. Для хранения данных используются базы данных (для них характерны большой объем и относительно небольшая стоимость информации), а для хранения знаний используется база знаний (небольшого объема, но исключительно дорогие информационные массивы).

Знаниями называют закономерности предметной области (принципы, связи, законы), которые были получены в результате практической деятельности и профессионального опыта [2].

#### **2.1.2 Стратегии получения знаний**

При формировании базы знаний ключевым вопрос является сам процесс получения знаний. Для названия этого процесса в литературе распространение получили несколько терминов: приобретение, извлечение и формирование знаний.

Приобретение знаний - это метод автоматизированного наполнения базы знаний посредством взаимодействия эксперта и специальной программы.

Под извлечение знаний понимается процесс взаимодействия инженера по знаниям с источником знаний (экспертом и др.). Процедура извлечения знаний длительна и трудоемка, так как инженеру по знаниям, вооруженными специальными знаниями по психологии, системному анализу, математической логике и пр., необходимо воссоздать модель предметной области, которой пользуются эксперты для принятия решения.

Формирование знаний — это термин, который закрепился в чрезвычайно перспективной и активно развивающейся области инженерии знаний, занимающаяся разработкой моделей, методов и алгоритмов обучения. Она включает индуктивные модели формирования знаний и автоматического порождения гипотез на основе обучающих выборок, обучение по аналогии и другие методы. Данные модели позволяют выявить причинно-следственные эмпирические зависимости в базах данных с неполной информацией, содержащих структурированные числовые и символьные объекты (часто в условиях неполноты информации) [6].

Формирование знаний – это процедура, которая заключается в анализа данных и выявлении скрытых закономерностей при помощи специального математического аппарата и программных средств.

Таким образом, можно выделить три основных способа получения знаний:

1. С использованием ЭВМ при наличии подходящего программного инструментария, другими словами - приобретение знаний;

2. С использованием программ обучения при наличии достаточно представительной выборки примеров принятия решений в предметной области и соответствующих пакетов прикладных программ, другими словами формирование знаний.

3. Без использования вычислительной техники за счет прямого взаимодействия инженера по знаниям и источника знаний (будь то эксперт, специальная литература или другие источники), другими словами - извлечение знаний;

### **2.1.3 Технология Data Mining**

Термин Data Mining зачастую переводится как добыча данных, раскопка данных или «извлечение зерен знаний из гор данных» [6]. Данную технологию очень точно определяет один из основателей этого направления Григорий Пиатецкий-Шапиро: «Data Mining – это процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности» [7].

Data Mining – это совокупность различных способов обнаружения знаний. Выбор метода определяется типом имеющихся данных и информацией, которую необходимо получить.

В перечень методов и алгоритмов Data Mining входят следующие: деревья решений, искусственные нейронные сети, метод опорных векторов, символьные правила, иерархические методы кластерного анализа, метод ограниченного перебора, методы ближайшего соседа и  $k$ -ближайшего соседа, байесовские сети, линейная регрессия, корреляционно - регрессионный анализ; неиерархические методы кластерного анализа, в том числе алгоритмы  $k$ -средних и  $k$ -медианы; методы поиска ассоциативных правил, эволюционное программирование и генетические алгоритмы, разнообразные методы визуализации данных и множество других методов.

После изучения выше перечисленных методов и алгоритмов, основываясь на опыт наших коллег из лаборатории информационных технологий в социальной сфере и медицине в сотрудничестве с городской больницей №3, для решения поставленных задач использован кластерный анализ и построение деревьев решений.

### **2.1.4. Кластерный анализ**

Кластеризацией называется группировка объектов основанных на данных, которые описывают сущность этих объектов. Объекты в рамках одного кластера должны быть «схожими» и отличаться от объектов, вошедших в другие кластеры. Точность

кластеризации зависит от того, насколько сильно объекты похожи внутри одного кластера и различны между объектами других кластеров.

Существуют различные методы кластерного анализа, их можно разделить на две группы:

1. Иерархические:

- Иерархические агломеративные методы (Agglomerative Nesting, AGNES);
- Иерархические дивизимные (делимые) методы (DIvisive ANAlysis, DIANA).

2. Неиерархические (итеративные) методы:

- Алгоритм  $k$ -средних ( $k$ -means);
- Алгоритм PAM (partitioning around Medoids).

Используя различные методы кластерного анализа, можем получать различные решения для одних и тех же данных.

В работе применён наиболее распространенный алгоритм  $k$ -средних. Алгоритм разбивает объекты на  $k$  групп, где каждая группа представляет один кластер. Основным отличием данного метода от иерархического кластер-анализа состоит в необходимости изначально определить число кластеров, на которое требуется разбить изучаемую совокупность. Поэтому, желательно до начала анализа иметь гипотезу о структуре исследуемой совокупности. С помощью данного метода была исследована структура медицинских данных, проводились случаи кластеризации для  $k = 4$ .

### 2.1.5. Построение деревьев решений

Деревья решений осуществляют разбиение пространства объектов в соответствии с некоторым набором правил разбиения. Эти правила являются логическими утверждениями в отношении той или иной переменной и могут быть истинными или ложными. Ключевыми здесь являются три обстоятельства: а) правила позволяют реализовать последовательную дихотомическую сегментацию данных, б) два объекта считаются похожими, если они оказываются в одном и том же сегменте разбиения, в) на каждом шаге разбиения увеличивается количество информации относительно исследуемой переменной (отклика).

Деревья решений являются одним из наиболее популярных методов решения многих практических задач, что обусловлено следующими причинами:

1. Деревья решений позволяют получать очень легко интерпретируемые модели, представляющие собой набор правил вида «если..., то...». Интерпретация облегчается, в том числе, за счет возможности представить эти правила в виде наглядной древовидной структуры.

2. В силу своего устройства деревья решений позволяют работать с переменными любого типа без необходимости какой-либо предварительной подготовки этих переменных для ввода в модель (например, логарифмирование, преобразование категориальных переменных в индикаторные, и т.п.).

3. Исследователю нет необходимости в явном виде задавать форму взаимосвязи между откликом и предикторами, как это, например, происходит в случае с обычными регрессионными моделями. Это оказывается особенно полезным при работе с данными большого объема, о свойствах которых мало что известно.

4. Деревья решений, по сути, автоматически выполняют отбор информативных предикторов и учитывают возможные взаимодействия между ними. Это, в частности, делает деревья решений полезным инструментом разведочного анализа данных.

5. Деревья решений можно эффективно применять к данным с пропущенными значениями, что очень полезно при решении практических задач, где наличие пропущенных значений – это, скорее, правило, чем исключение.

6. Деревья решений одинаково хорошо применимы как к количественным, так и к качественным зависимым переменным.

## **2.2. Выбор инструмента решения**

Продуктивность выполняемой работы зависит от используемых инструментов. Применение статистического программного обеспечения существенно повышает быстроту и эффективность обработки данных.

Все научные математические приложения делятся на две группы: программы с графическим интерфейсом и без него. Надо отметить, что графический интерфейс не влияет на качество программного продукта. Но тем не менее, подобное разделение имеет огромное практическое значение. На данный момент насчитываются десятки качественных статистических пакетов, рассмотрим наиболее популярные из них [8].

1. Система SAS, зародившаяся в 60-х годах XX века в Университете Северной Каролины как приложение для анализа результатов сельскохозяйственных исследований. В наше время данная система продолжает активно развиваться в компании SAS Institute, которая выпустила уже 9 версию этой программы. SAS используется в различных областях: самые разнообразные научные исследования, бизнес аналитика и т. д.

В состав системы входят модули, которые выполняют определенный круг задач. В системе SAS реализован собственный язык программирования. SAS обладает возможностью загрузки данных из внешних файлов и вводом их непосредственно в окно терминала. Данная система позволяет проводить статистическую обработку данных

разного уровня сложности. В программе реализован консольный режим и графический интерфейс, представляющий собой графическую оболочку для упрощенного ввода команд языка программирования SAS.

2. Язык программирования R. Он основан на языке программирования S. Система R обладает широкими возможностями для статистической обработки данных и графического отображения, а оконный интерфейс можно установить как дополнительное приложение. Синтаксис языка достаточно прост и понятен в изучении.

3. Stata, система американской корпорации StataCorp. Ввод данных предусмотрен как путем загрузки из внешних файлов, так и с использованием встроенного табличного редактора. Принципы работы с приложением Stata не отличаются от таковых при использовании описанных выше программ.

4. MedCalc, разрабатываемая с 1993 года бельгийской компанией MedCalc Software. Данная система позиционируется как полноценное статистическое приложение, созданное в соответствии с потребностями исследователей биомедицинского направления. Разработчики акцентируют внимание исследователей на удобстве использования MedCalc для анализа ROC-кривых. MedCalc адаптирован только под одну операционную систему - Windows.

5. Statistica, разрабатываемая компанией StatSoft. Программа обладает хорошим графическим интерфейсом, а также имеет встроенный язык программирования. Следует отметить, что почти безграничные возможности в статобработке, предоставляемые данными инструментами, требуют от компьютера больших ресурсов. Statistica разработана только под Windows, что является одним из ее недостатков.

6. Программа SPSS, название которой - это аббревиатура от Statistical Package for the Social Sciences, относительно недавно стала принадлежать компании IBM и сменила название на PASW (Predictive Analytics SoftWare). В системе реализованы широкие графические возможности. Она состоит из огромного количества разнообразных категорий и типов графиков, в том числе научные, деловые, трехмерные и двухмерные графики в различных системах координат, специализированные статистические графики — гистограммы, матричные, категоризованные графики и др. Для работы SPSS необходимо не менее 1 Гб оперативной памяти.

Проведем сравнительный анализ данных систем на наибольшую эффективность в задачах статистического анализа (таблица 1).

Таблица 1 - Сравнительный анализ инструментов решений

| Признак/Инструмент              | SAS | R | Stata | MedCalc | Statistica | SPSS |
|---------------------------------|-----|---|-------|---------|------------|------|
| Наличие графического интерфейса | +   | + | +     | +       | +          | +    |
| Наличие консольного ввода       | +   | + | +     | +       | +          | +    |
| Свободное распространение       | -   | + | -     | -       | -          | -    |
| Работа с графикой               | +   | + | +     | +       | +          | +    |
| Поддержка ОС Linux и Windows    | +   | + | +     | -       | -          | +    |
| Простота в понимании и изучении | +   | + | +     | +       | -          | -    |
| Ресурсоэффективная              | +   | + | +     | +       | -          | -    |

На основе проведенного сравнительного анализа для решения поставленной задачи был выбран язык R. В качестве среды разработки выбрано – Rstudio. RStudio — свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R, который предназначен для статистической обработки данных и работы с графикой.

### 2.2.1. Скриптовый язык R

R — это язык программирования, который разработан для применения в таких областях, как разведочный анализ данных, классические статистические тесты и высокоуровневая графика. R основан на языке программирования S, который был разработан в Bell Labs, и является его альтернативной реализацией. Язык R был разработан сотрудниками статистического факультета Оклендского университета Россом Айхэкой и Робертом Джентлменом (первая буква их имён — R); язык и среда поддерживаются и развиваются организацией R Foundation [11].

Язык R обладает обширной и постоянно расширяющейся библиотекой пакетов, что делает его одним из лидеров среди систем для статистического анализа и добычи данных.

### 2.2.2. Типы данных языка R

Объекты данных в R можно делить на следующие классы (т.е. типы объектов):

- ° numeric – объекты, к которым относятся целочисленные (integer) и действительные числа (double);
- ° logical – логические объекты, которые принимают только два значения: FALSE (сокращенно F) и TRUE (T);
- ° character – символьные объекты (значения переменных задаются в двойных, либо одинарных кавычках).

Таблица данных (data frame) представляет собой объект R, по структуре напоминающий лист электронной таблицы Microsoft Excel. Каждый столбец таблицы представляет собой вектор, содержащий данные определенного типа. Таблицы данных – это основной класс объектов R, используемых для хранения данных. В основном таблицы подготавливаются при помощи внешних приложений и после чего загружаются в среду R

### **2.2.3. Импортирование данных в R**

Правила подготовки загружаемых файлов:

- Таблица с данными не должно содержать пустых ячеек. В случае отсутствия значения вместо них необходимо ввести NA.

- Импортируемая таблица с данными должна быть преобразована в простой текстовый файл с одним из допустимых расширений. На практике обычно используются файлы с расширением .txt и .csv.

- Первая строка в импортируемой таблице должна содержать заголовки столбцов-переменных. Такая строка – удобный, но не обязательный элемент загружаемого файла.

Имена столбцов таблицы лучше присвоить с соблюдением правил идентификации переменных R, т.е. исключить пробелы и другие специальные символы, кроме точки и подчеркивания. Во избежание проблем, связанных с кодировкой, все текстовые величины в импортируемых файлах стоит создавать с использованием букв латинского алфавита.

## Глава 3 Решение поставленной задачи и описание полученных результатов

### 3.1. Подготовка данных к анализу

Критически важным является подготовка исходных показателей, так как их применение способствует успешному выполнению задачи анализа данных. Исходные данные находятся в таблице Excel (Приложение А).

Исходные данные были преобразованы к табличному виду и отредактированы. Вся таблица исходных данных соответствует одному типу переменных. Числовые значения всех медицинских показателей, входящих в матрицу, относятся к количественному типу переменных.

### 3.2. Алгоритм выявления скрытых закономерностей

С целью обработки клинико-лабораторных показателей, построено дерево решений, выполнен кластерный анализ с применением пакетов языка R. На основе спроектированной в рамках гранта РФФИ системы поддержки медицинских исследований был разработан алгоритм выявления скрытых закономерностей (рисунок 3.1).



Рисунок 3.1 – Схема алгоритма выявления скрытых закономерностей



```

1 100.59385 110.23231 89.31000 92.98615 95.31923 105.45385 77.62308 86.43077 74.85
615
2 78.54375 85.49313 87.87500 89.59063 73.47250 78.34000 71.13750 77.85625 68.25
000
3 54.42739 59.67391 68.91217 74.01261 44.25174 45.48087 37.81913 41.15957 35.81
478
4 70.26129 78.30323 79.21258 85.89935 71.37742 76.21290 55.84419 60.72968 53.66
452
      MOC25_2 MOC50_1 MOC50_2 MOC75_1 MOC75_2      cdyn1      cdyn2      Cstat1      Cst
at2
1 82.60769 65.37692 76.36308 58.76923 68.54615 0.12046154 0.12738462 0.1520769 0.15469
23
2 78.15000 61.95000 68.64375 56.18125 62.76250 0.15075000 0.15912500 0.1770000 0.18187
50
3 38.48826 30.19826 31.45174 25.95957 26.56609 0.09047826 0.08730435 0.1325652 0.11782
61
4 58.51290 47.31935 51.01613 42.34839 44.36774 0.10593548 0.10990323 0.1351613 0.13596
77
      RVд1      RVд2      RVыд1      RVыд2      Wобщ1      Wобщ2      Wуд1      Wуд2      Wн.
эл1
1 3.564615 2.837692 4.277692 3.413846 0.4357692 0.3058462 0.05230769 0.04061538 0.3740
769
2 3.838125 3.114375 4.910625 3.815000 0.3656250 0.3077500 0.04200000 0.03800000 0.2856
250
3 5.582609 4.443478 7.395652 5.715652 0.6857174 0.4971304 0.08182609 0.06786957 0.6210
000
4 4.750000 3.998065 6.494194 4.921613 0.5502581 0.4406129 0.06735484 0.05967742 0.4912
581
      Wн.эл2      Wэл1      Wэл2 W МОД10_1 W МОД10_2 W МОД15_1 W МОД15_2 W МВЛобщ1 WМВЛ
общ2
1 0.2652308 0.2171538 0.1583077 0.5033846 0.3700000 0.9423077 0.6584615 22.44423 25.1
3462
2 0.2686250 0.2141875 0.1768750 0.4361250 0.3860625 0.7439375 0.6343750 17.89494 21.0
5625
3 0.4587826 0.3151304 0.2685217 0.8122174 0.6238696 1.3847826 1.1151739 13.52570 14.3
7026
4 0.3949032 0.2631935 0.2412581 0.6443871 0.5509677 1.0401290 0.8609677 24.57839 27.6
5090
      WМВЛуд1      WМВЛуд2
1 0.2256154 0.2366923
2 0.2363125 0.2841875
3 0.3094783 0.3167826
4 0.3410323 0.3654516

Clustering vector:
[1] 2 4 2 2 2 3 4 3 1 4 4 3 4 2 3 3 4 1 3 4 4 4 4 3 3 4 1 3 4 3 4 3 4 4 1 4 4 3 4 3
3 3 3
[45] 4 4 4 3 4 2 1 3 1 4 4 3 3 4 1 4 4 2 4 3 4 4 3 2 3 2 2 1 2 2 2 2 1 1 1 2 1 1
2

within cluster sum of squares by cluster:
[1] 45341.97 28495.67 64792.54 74257.86
(between_SS / total_SS = 60.3 %)

```

Рисунок 3.2 – Результат работы скрипта

Чтобы убедиться в результатах анализа, определим средние значения всех анализируемых параметров в каждом из кластеров:

```
> aggregate(data,by=list(clus$cluster), FUN=mean),
```

где `aggregate(x,by,FUN)` — функция, которая разбивает таблицу данных `x` на

отдельные наборы данных, применяет к этим наборам определенную функцию FUN и возвращает результат в удобном для чтения формате.

Результат работы скрипта представлен на рисунке 3.3.

| Group.1  | Возраст    | Рост      | Вес        | МОД1       | МОД2      | ЖЕЛ1      | ЖЕЛ2       |              |
|----------|------------|-----------|------------|------------|-----------|-----------|------------|--------------|
| ФЖЕЛ1    |            |           |            |            |           |           |            |              |
| 1        | 1          | 31.38462  | 169.8462   | 71.00000   | 8.139231  | 7.326923  | 104.60000  | 112.06154 10 |
| 6.28923  |            |           |            |            |           |           |            |              |
| 2        | 2          | 31.43750  | 168.6875   | 63.12500   | 8.759375  | 8.262500  | 86.97625   | 93.35000 8   |
| 1.35125  |            |           |            |            |           |           |            |              |
| 3        | 3          | 37.78261  | 166.3478   | 63.69565   | 8.400000  | 7.197826  | 67.81435   | 70.14652 6   |
| 0.17348  |            |           |            |            |           |           |            |              |
| 4        | 4          | 39.61290  | 164.4839   | 67.41935   | 8.199032  | 7.514839  | 85.93871   | 87.46032 7   |
| 7.75161  |            |           |            |            |           |           |            |              |
| ФЖЕЛ2    | ОФВ1_1     | ОФВ1_2    | ОФВ1/ЖЕП 1 | ОФВ1/ЖЕЛ 2 | МВЛ1      | МВЛ2      | ПОС1       | ПОС2         |
| 1        | 114.16923  | 100.5938  | 110.2323   | 89.3100    | 92.98615  | 95.31923  | 105.45385  | 77.6230      |
| 8        | 86.43077   |           |            |            |           |           |            |              |
| 2        | 86.19312   | 78.5437   | 85.4931    | 87.8750    | 89.59063  | 73.47250  | 78.34000   | 71.1375      |
| 0        | 77.85625   |           |            |            |           |           |            |              |
| 3        | 65.01826   | 54.4273   | 59.6739    | 68.9121    | 74.01261  | 44.25174  | 45.48087   | 37.8191      |
| 3        | 41.15957   |           |            |            |           |           |            |              |
| 4        | 82.38710   | 70.2619   | 78.3032    | 79.2125    | 85.89935  | 71.37742  | 76.21290   | 55.8441      |
| 9        | 60.72968   |           |            |            |           |           |            |              |
| МОС25_1  | МОС25_2    | МОС50_1   | МОС50_2    | МОС75_1    | МОС75_2   | Cdyn1     | Cdyn2      | Cstat1       |
| 1        | 74.85615   | 82.60769  | 65.37692   | 76.36308   | 58.76923  | 68.54615  | 0.12046154 | 0.12738462   |
| 0        | 0.1520769  |           |            |            |           |           |            |              |
| 2        | 68.25000   | 78.15000  | 61.95000   | 68.64375   | 56.18125  | 62.76250  | 0.15075000 | 0.15912500   |
| 0        | 0.1770000  |           |            |            |           |           |            |              |
| 3        | 35.81478   | 38.48826  | 30.19826   | 31.45174   | 25.95957  | 26.56609  | 0.09047826 | 0.08730435   |
| 0        | 0.1325652  |           |            |            |           |           |            |              |
| 4        | 53.66452   | 58.51290  | 47.31935   | 51.01613   | 42.34839  | 44.36774  | 0.10593548 | 0.10990323   |
| 0        | 0.1351613  |           |            |            |           |           |            |              |
| Cstat2   | РВД1       | РВД2      | РВЫД1      | РВЫД2      | Вообщ1    | Вообщ2    | Вуд1       | Вуд2         |
| 1        | 0.1546923  | 3.564615  | 2.837692   | 4.277692   | 3.413846  | 0.4357692 | 0.3058462  | 0.0523076    |
| 0        | 0.0406153  |           |            |            |           |           |            |              |
| 2        | 0.1818750  | 3.838125  | 3.114375   | 4.910625   | 3.815000  | 0.3656250 | 0.3077500  | 0.0420000    |
| 0        | 0.0380000  |           |            |            |           |           |            |              |
| 3        | 0.1178261  | 5.582609  | 4.443478   | 7.395652   | 5.715652  | 0.6857174 | 0.4971304  | 0.0818260    |
| 0        | 0.0678695  |           |            |            |           |           |            |              |
| 4        | 0.1359677  | 4.750000  | 3.998065   | 6.494194   | 4.921613  | 0.5502581 | 0.4406129  | 0.067354     |
| 0        | 0.05967742 |           |            |            |           |           |            |              |
| ВН.эл1   | ВН.эл2     | Вэл1      | Вэл2       | W МОД10_1  | W МОД10_2 | W МОД15_1 | W МОД15_2  | W МВЛобщ1    |
| 1        | 0.3740769  | 0.26523   | 0.2171538  | 0.1583077  | 0.5033846 | 0.37000   | 0.9423077  | 0.6584615    |
| 22       | 44423      |           |            |            |           |           |            |              |
| 2        | 0.2856250  | 0.26862   | 0.2141875  | 0.1768750  | 0.4361250 | 0.38606   | 0.7439375  | 0.6343750    |
| 17       | 89494      |           |            |            |           |           |            |              |
| 3        | 0.6210000  | 0.45878   | 0.3151304  | 0.2685217  | 0.8122174 | 0.62389   | 1.3847826  | 1.1151739    |
| 13       | 52570      |           |            |            |           |           |            |              |
| 4        | 0.4912581  | 0.39490   | 0.2631935  | 0.2412581  | 0.6443871 | 0.55096   | 1.0401290  | 0.8609677    |
| 24       | 57839      |           |            |            |           |           |            |              |
| WMBЛобщ2 | WMBЛуд1    | WMBЛуд2   |            |            |           |           |            |              |
| 1        | 25.13462   | 0.2256154 | 0.2366923  |            |           |           |            |              |
| 2        | 21.05625   | 0.2363125 | 0.2841875  |            |           |           |            |              |
| 3        | 14.37026   | 0.3094783 | 0.3167826  |            |           |           |            |              |
| 4        | 27.65090   | 0.3410323 | 0.3654516  |            |           |           |            |              |

Рисунок 3.3 – Результат работы скрипта

На рисунках 3.4-3.6 изображены графики средних значения в кластерах.

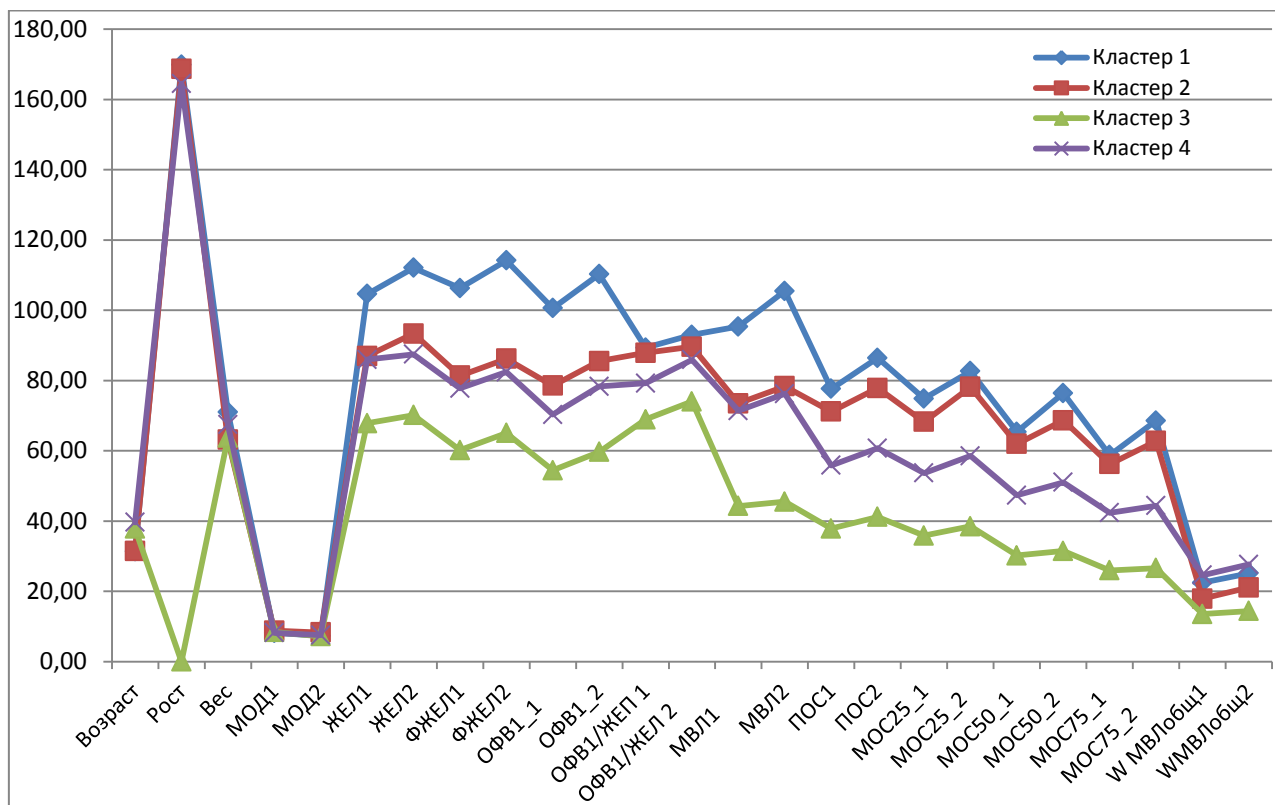


Рисунок 3.4 – График средних значения в кластерах

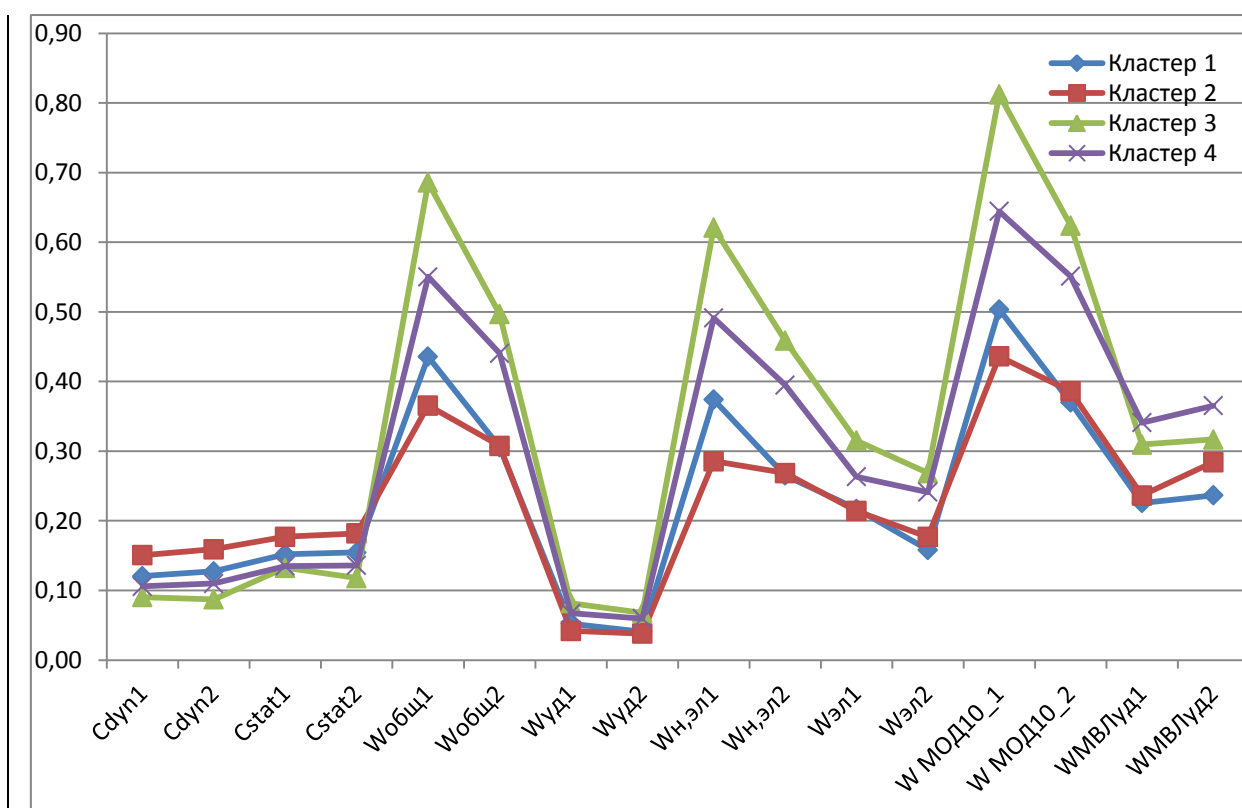


Рисунок 3.5 – График средних значения в кластерах (продолжение)

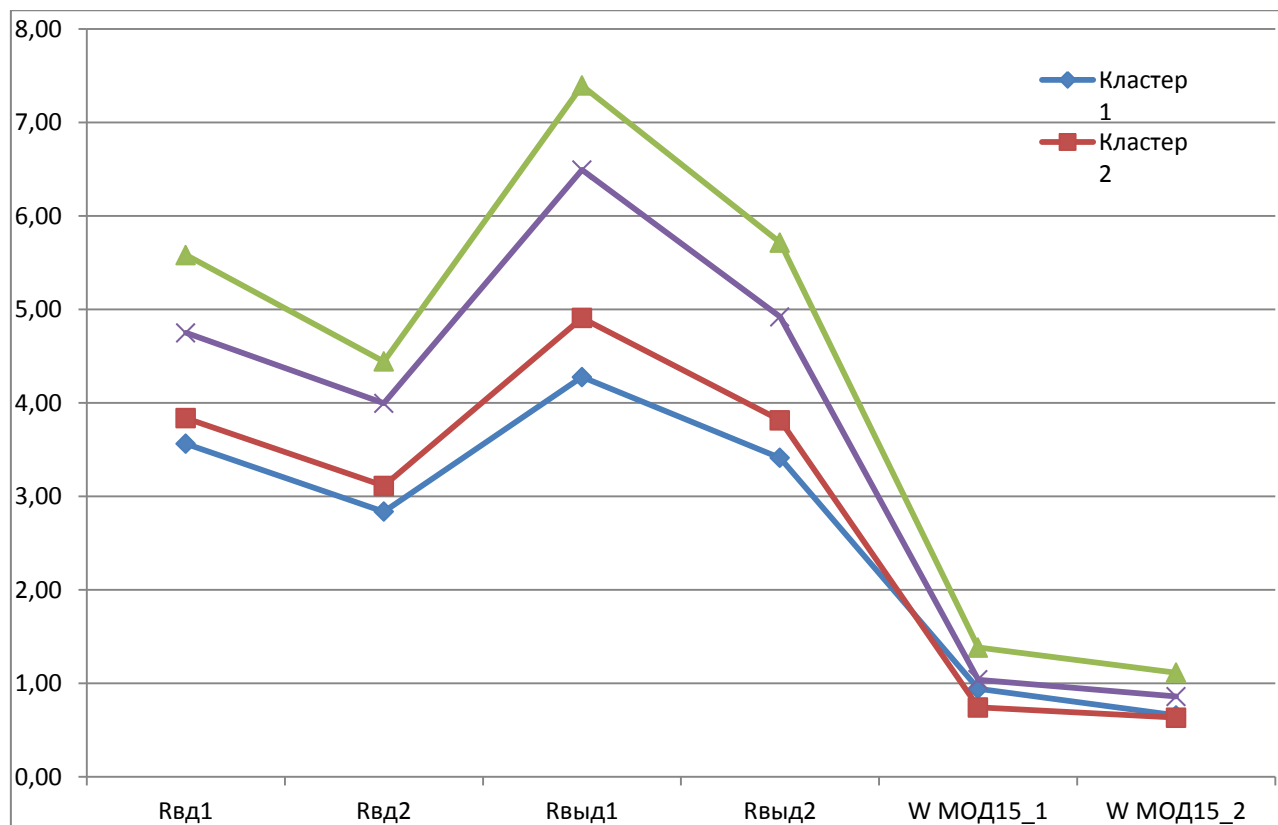


Рисунок 3.6 – График средних значения в кластерах (продолжение)

Из полученной таблицы различие между записями в разных кластерах уже видно. Теперь присвоим номера кластеров каждой из записей исходного набора данных:

```
> data <- data.frame(data, clus$cluster)
```

С помощью библиотеки `cluster` можно построить более типичный график для кластерного анализа (рисунок 3.7):

```
> library(cluster)
```

```
> clusplot(data, clus$cluster, color=TRUE, shade=TRUE, labels=2, lines=0)
```

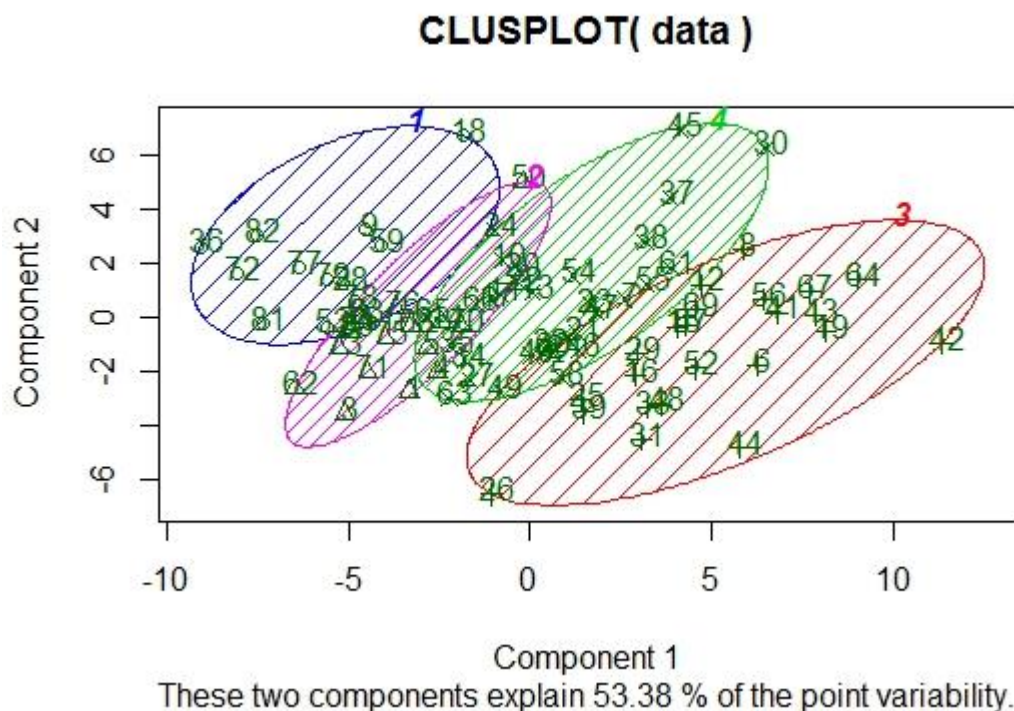


Рисунок 3.7 - График кластерного анализ

Анализ достоверности полученных данных приведен в таблице 2.

Таблица 2 - Анализ достоверности полученных данных

| Номер кластера   | Число достоверных данных | Число недостоверных данных | Процент достоверности данных, % |
|------------------|--------------------------|----------------------------|---------------------------------|
| Кластер 1 - БАПИ | 26                       | 3                          | 89,65                           |
| Кластер 2 - БАНП | 15                       | 2                          | 88,23                           |
| Кластер 3 - БАСП | 24                       | 3                          | 88,89                           |
| Кластер 4 - ПО   | 8                        | 1                          | 80                              |

### 3.4. Построение деревьев решений

В общем случае может быть использовано несколько алгоритмов построения деревьев на основе различных схем и критериев оптимизации. Функция *rpart()* из одноименного пакета выполняет рекурсивный выбор для каждого следующего узла таких разделяющих значений, которые приводят к минимальной сумме квадратов внутригрупповых отклонений  $D_t$  для всех  $t$  узлов дерева. Для оценки качества построенного дерева  $T$  в ходе его оптимизации используется следующая совокупность критериев:

- штраф за сложность модели (*cost complexity*), включающий штрафной множитель за каждую неотсечённую ветвь  $T = \sum t D_t + \lambda_t$ ;
- девианс  $D_0$  для нулевого дерева (т.е. оценка изменчивости в исходных данных);

- относительный параметр стоимости сложности  $C_p = \lambda/D_0$ ;
- относительная ошибка обучения для дерева из  $t$  узлов  $RELe = \sum t D_t / D_0$ ;
- ошибка перекрестной проверки ( $CVer$ ) с разбиением на 10 блоков, также отнесенная к девиансу нуль-дерева  $D_0$ ;  $CVer$ , как правило, больше, чем  $RELe$ ;
- стандартное отклонение ( $SE$ ) ошибки перекрестной проверки.

Лучшим считается дерево, состоящее из такого количества ветвей  $t$ , для которого сумма ( $CVer + SECVe + SE$ ) является минимальной.

Подключим библиотеку «*rpart*» и построим дерево на основе имеющихся данных:

```
> library(rpart)
> (rt.a1 <- rpart(a1 ~ ., data = algae[, 1:12]))
```

Результат работы скрипта скрипта представлен на рисунке 3.8.

```
n= 83
node), split, n, loss, yval, (yprob)
      * denotes terminal node
 1) root 83 54 BANP (0.34939759 0.28915663 0.21686747 0.14457831)
   2) ЖЕЛ2< 111.225 72 45 BANP (0.37500000 0.33333333 0.23611111 0.05555
556)
     4) ОФВ1_2< 82.55 51 27 BANP (0.47058824 0.21568627 0.25490196 0.058
82353)
       8) ЖЕЛ1>=63.67 41 18 BANP (0.56097561 0.12195122 0.24390244 0.073
17073)
         16) МОД1>=8.225 18 6 BANP (0.66666667 0.16666667 0.00000000 0.1
6666667) *
           17) МОД1< 8.225 23 12 BANP (0.47826087 0.08695652 0.43478261 0.0
0000000)
             34) ФЖЕЛ2< 76.95 11 3 BANP (0.72727273 0.09090909 0.18181818
0.00000000) *
               35) ФЖЕЛ2>=76.95 12 4 BASP (0.25000000 0.08333333 0.66666667
0.00000000) *
                 9) ЖЕЛ1< 63.67 10 4 BAPI (0.10000000 0.60000000 0.30000000 0.000
00000) *
                   5) ОФВ1_2>=82.55 21 8 BAPI (0.14285714 0.61904762 0.19047619 0.047
61905) *
                     3) ЖЕЛ2>=111.225 11 3 PD (0.18181818 0.00000000 0.09090909 0.7272727
3) *
```

Рисунок 3.8 – Дерево решений

Приведенной командой мы построили полное дерево без обрезания ветвей. В каждой строке представлены по порядку: условие разделения, число наблюдений, соответствующих этому условию, девианс (в данном случае - это эквивалент суммы квадратов отклонений от группового среднего) и среднее значение отклика для выделенной ветви. Например, перед первой итерацией общее множество из 83 наблюдений имеет среднее значение  $m = \text{BANP}$  при девиансе  $D = 54$ . При  $\text{ЖЕЛ2} < 111,225$  это множество делится на две части: 2) 72 наблюдения ( $m = \text{BANP}$ ,  $D = 45$ ) и 3) 11 наблюдений ( $m = \text{PD}$ ,  $D = 3$ ). Дальнейшие разбиения каждой из этих двух частей аналогичны.

Представим данное дерево графически. Для это воспользуемся функцией `prp()` из пакета `part.plot`.

```
> prp(rt.a1)
```

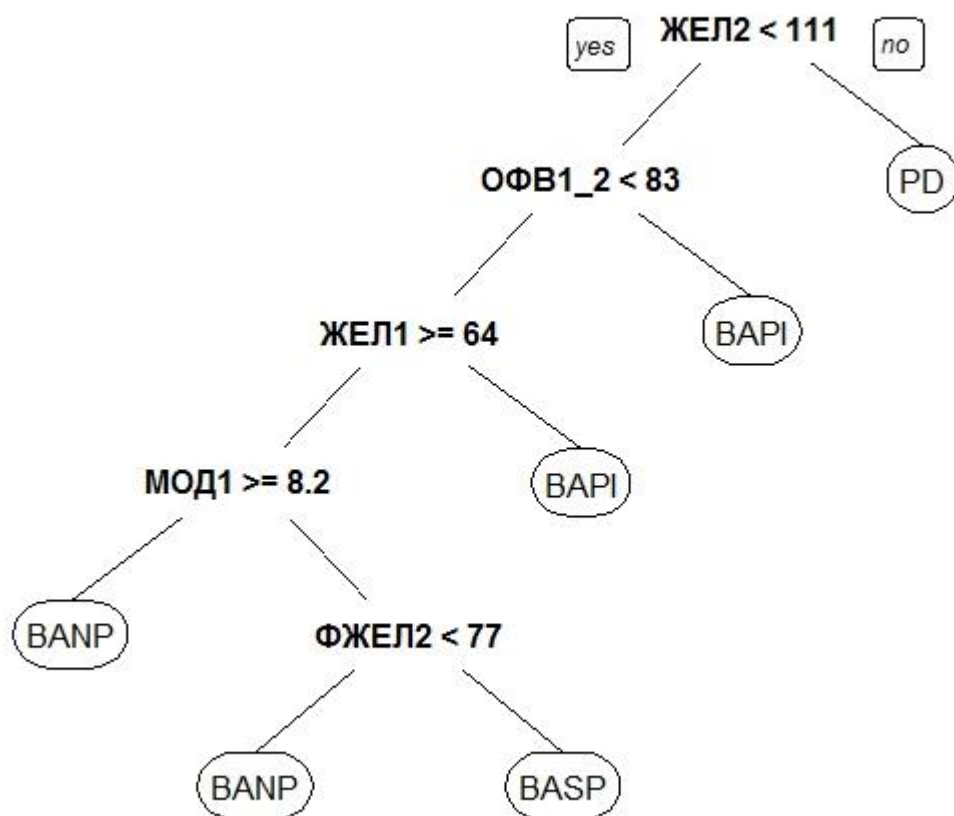


Рисунок 3.9 - Графическое отображение дерева решений

Данное дерево дает нам следующие логические правила:

1. Если  $\text{ЖЕЛ2} \geq 111.225$ , то группа БА = PD.
2. Если  $\text{ЖЕЛ2} < 111.225$  и  $\text{ОФВ1\_2} \geq 82.55$ , то группа БА = ВАРІ.
3. Если  $\text{ЖЕЛ2} < 111.225$  и  $\text{ОФВ1\_2} < 82.55$  и  $\text{ЖЕЛ1} < 63.67$ , то группа БА = ВАРІ.
4. Если  $\text{ЖЕЛ2} < 111.225$  и  $\text{ОФВ1\_2} < 82.55$  и  $\text{ЖЕЛ1} \geq 63.67$  и  $\text{МОД1} \geq 8.225$ , то группа БА = ВАНР.
5. Если  $\text{ЖЕЛ2} < 111.225$  и  $\text{ОФВ1\_2} < 82.55$  и  $\text{ЖЕЛ1} \geq 63.67$  и  $\text{МОД1} < 8.225$  и  $\text{ФЖЕЛ2} \geq 76.95$ , то группа БА = ВАСР.
6. Если  $\text{ЖЕЛ2} < 111.225$  и  $\text{ОФВ1\_2} < 82.55$  и  $\text{ЖЕЛ1} \geq 63.67$  и  $\text{МОД1} < 8.225$  и  $\text{ФЖЕЛ2} < 76.95$ , то группа БА = ВАНР.

Анализ достоверности полученных данных приведен в таблице 3.

Таблица 3 - Анализ достоверности полученных данных

| Логическое правило                                                                                                    | Число достоверных данных | Число недостоверных данных | Процент достоверности данных, % |
|-----------------------------------------------------------------------------------------------------------------------|--------------------------|----------------------------|---------------------------------|
| 1. Если ЖЕЛ2 $\geq$ 111.225, то группа БА = PD.                                                                       | 9                        | 2                          | 81,81                           |
| 2. Если ЖЕЛ2 < 111.225 и ОФВ1_2 $\geq$ 82.55, то группа БА = VARI.                                                    | 15                       | 3                          | 83,33                           |
| 3. Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 < 63.67, то группа БА = VARI.                                          | 8                        | 1                          | 80                              |
| 4. Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 $\geq$ 63.67 и МОД1 $\geq$ 8.225, то группа БА = BANP.                 | 16                       | 2                          | 88,89                           |
| 5. Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 $\geq$ 63.67 и МОД1 < 8.225 и ФЖЕЛ2 $\geq$ 76.95, то группа БА = BASP. | 10                       | 2                          | 83,33                           |
| 6. Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 $\geq$ 63.67 и МОД1 < 8.225 и ФЖЕЛ2 < 76.95, то группа БА = BANP.      | 13                       | 3                          | 81,25                           |

Как видно из таблицы 3, результаты построения дерева решений имеют достаточно высокий показатель достоверности данных.

Коллектив исследователей, работающих в рамках гранта РФФИ 14-07-00675 получал продуктивные модели при помощи пакета WizWhy. Система WizWhy предприятия WizSoft является программным средством анализа данных, поиска логических правил в данных. Данная система на сегодняшний день является одним из лидеров на рынке продуктов Data Mining. Но у данного подхода есть существенный недостаток – слишком большой объем логических правил. В данном случае было найдено более 1000 правил.

Пакет языка R позволяет отобрать наиболее значимые (исходя из высокого процента достоверности данных) логические правила.

В таблице 4 представлено соотнесение правил, полученных при помощи пакета языка R и системой WizWhy.

Таблица 4 — Соотнесение логических правил

| Язык R                                       | Система WizWhy                                                                                                                                                                                                                                                   |
|----------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Если ЖЕЛ2 $\geq$ 111.225, то группа БА = PD. | <p>If ЖЕЛ2 is <b>111,23 ... 144,81</b> (average = <b>128,02</b>)</p> <p>Then</p> <p><b>Group is <u>PD</u></b></p> <p>Rule's probability: <b>0,82</b></p> <p>The rule exists in <b>11</b> records.</p> <p>Significance Level: Error probability &lt; 0,000001</p> |

|                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                            | <p><i>Positive Examples (records' serial numbers):</i><br/> <b>73, 74, 75, 76, 77, 78, 80, 81, 82</b></p> <p><i>Negative Examples (records' serial numbers):</i><br/> <b>74, 83</b></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Если ЖЕЛ2 < 111.225 и ОФВ1_2 >= 82.55, то группа БА = ВАРІ.                                | <p>If ЖЕЛ2 is <u>48,6 ... 110,23</u> (average = <u>79,9</u>)<br/> and ОФВ1_2 is <u>82,55 ... 159,38</u> (average = <u>120,96</u>)</p> <p>Then<br/> <b>Group is ВАРІ</b><br/> Rule's probability: <b>0,833</b><br/> The rule exists in <b>18</b> records.<br/> Significance Level: Error probability &lt; 0,000001</p> <p><i>Positive Examples (records' serial numbers):</i><br/> <b>3, 5, 6, 7, 8, 10, 12, 13, 15, 16, 17, 18, 20, 21, 23</b></p> <p><i>Negative Examples (records' serial numbers):</i><br/> <b>1, 2, 4</b></p>                                                                                                                                       |
| Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 < 63.67, то группа БА = ВАРІ.                  | <p>If ЖЕЛ2 is <u>48,6 ... 110,23</u> (average = <u>79,9</u>)<br/> and ОФВ1_2 is <u>38,3 ... 82,63</u> (average = <u>60,46</u>)</p> <p>and ЖЕЛ1 is <u>49,7 ... 61,65</u> (average = <u>55,67</u>)</p> <p>Then<br/> <b>Group is ВАРІ</b><br/> Rule's probability: <b>0,8</b><br/> The rule exists in <b>9</b> records.<br/> Significance Level: Error probability &lt; 0,000001</p> <p><i>Positive Examples (records' serial numbers):</i><br/> <b>4, 11, 21, 22, 25, 26, 29, 27</b></p> <p><i>Negative Examples (records' serial numbers):</i><br/> <b>3</b></p>                                                                                                         |
| Если ЖЕЛ2 < 111.225 и ОФВ1_2 < 82.55 и ЖЕЛ1 >= 63.67 и МОД1 >= 8.225, то группа БА = ВАНР. | <p>If ЖЕЛ2 is <u>48,6 ... 110,23</u> (average = <u>79,9</u>)<br/> and ОФВ1_2 is <u>38,3 ... 82,63</u> (average = <u>60,46</u>)</p> <p>and ЖЕЛ1 is <u>63,67 ... 131,65</u> (average = <u>97,66</u>)<br/> and МОД1 is <u>8,23 ... 12</u> (average = <u>10,11</u>)</p> <p>Then<br/> <b>Group is ВАНР</b><br/> Rule's probability: <b>0,889</b><br/> The rule exists in <b>18</b> records.<br/> Significance Level: Error probability &lt; 0,000001</p> <p><i>Positive Examples (records' serial numbers):</i><br/> <b>27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 38, 40, 41, 42, 43, 45</b></p> <p><i>Negative Examples (records' serial numbers):</i><br/> <b>36, 39</b></p> |
| Если ЖЕЛ2 < 111.225 и ОФВ1_2 <                                                             | <p>If ЖЕЛ2 is <u>48,6 ... 110,23</u> (average = <u>79,9</u>)</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

|                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>82.55 и ЖЕЛ1 <math>\geq</math> 63.67 и МОД1 &lt; 8.225 и ФЖЕЛ2 <math>\geq</math> 76.95, то группа БА = BASP.</p>                         | <p>and <b>ОФВ1_2</b> is <u>38,3 ... 82,63</u> (average = <u>60,46</u>)</p> <p>and <b>ЖЕЛ1</b> is <u>63,67 ... 131,65</u> (average = <u>97,66</u>)</p> <p>and <b>МОД1</b> is <u>4,64 ... 8,1</u> (average = <u>6,37</u>)</p> <p>and <b>ФЖЕЛ2</b> is <u>76,95 ... 144,5</u> (average = <u>110,72</u>)</p> <p>Then</p> <p><b>Group</b> is <u>BASP</u></p> <p>Rule's probability: <b>0,833</b></p> <p>The rule exists in <b>12</b> records.</p> <p>Significance Level: Error probability &lt; 0,000001</p> <p>Positive Examples (records' serial numbers): <b>43, 44, 45, 46, 47, 48, 49, 51, 52, 53</b></p> <p>Negative Examples (records' serial numbers): <b>42, 50</b></p>                                                                                                |
| <p>Если ЖЕЛ2 &lt; 111.225 и ОФВ1_2 &lt; 82.55 и ЖЕЛ1 <math>\geq</math> 63.67 и МОД1 &lt; 8.225 и ФЖЕЛ2 &lt; 76.95, то группа БА = BANP.</p> | <p>If <b>ЖЕЛ2</b> is <u>48,6 ... 110,23</u> (average = <u>79,9</u>)</p> <p>and <b>ОФВ1_2</b> is <u>38,3 ... 82,63</u> (average = <u>60,46</u>)</p> <p>and <b>ЖЕЛ1</b> is <u>63,67 ... 131,65</u> (average = <u>97,66</u>)</p> <p>and <b>МОД1</b> is <u>4,64 ... 8,1</u> (average = <u>6,37</u>)</p> <p>and <b>ФЖЕЛ2</b> is <u>36,1 ... 76,75</u> (average = <u>56,42</u>)</p> <p>)</p> <p>Then</p> <p><b>Group</b> is <u>BANP</u></p> <p>Rule's probability: <b>0,812</b></p> <p>The rule exists in <b>16</b> records.</p> <p>Significance Level: Error probability &lt; 0,000001</p> <p>Positive Examples (records' serial numbers): <b>55, 56, 57, 58, 59, 61, 62, 63, 64, 66, 67, 68, 69</b></p> <p>Negative Examples (records' serial numbers): <b>54, 60, 65</b></p> |

Согласно таблице 4, найденные логические правила средствами языка R содержатся в продукционных моделях системы WizWhy.

### 3.5. Анализ полученных результатов

В ходе выполнения разработанного в рамках ВКР алгоритма были произведены кластерный анализ и построение дерева решений.

Данные были поделены на 4 кластера, соответствующих 4-м группам БА. Так как достоверность этих результатов была выше 80%, то можно сделать вывод, что предложенная томскими медиками (Языков К.Г., Немеров Е.В.) психосинергетическая модель психогенной бронхиальной астмы имеет право на существование.

А построение дерева решений позволило получить логические правила, по которым делятся данные на группы БА.

## Глава 4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

### 4.1. Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения

#### 4.1.1. Потенциальные потребители результатов исследования

Для анализа потребителей результатов исследования необходимо рассмотреть целевой рынок и провести его сегментирование.

Целевой рынок проекта – медицинские учреждения. Критерии сегментирования данной работы - профессия в медицинской сфере, социальная принадлежность, ступени образования, научная степень, звание, вид интернет-ресурса. Выделим наиболее значимые для разрабатываемой информационной системы критерии: вид интернет-ресурса и профессия в медицинской сфере. На основе выявленных данных построим карту сегментирования, она представлена на рисунке 4.1.

|           |                              | Вид интернет-ресурса |                  |                       |
|-----------|------------------------------|----------------------|------------------|-----------------------|
|           |                              | Корпоративный сайт   | Веб - приложение | Информационный портал |
| Профессия | Врач общей                   |                      |                  |                       |
|           | Научный сотрудник            |                      |                  |                       |
|           | Главный врач мед. учреждения |                      |                  |                       |

Рисунок 4.1- Карта сегментирования рынка услуг по разработке интернет-ресурсов:



Информационная система «Здоровье»



Информационная система «Анализ данных»

В результате построения карты сегментирования выявлено, какие ниши на рынке услуг по разработке интернет-ресурсов не заняты конкурентами или где уровень конкуренции низок.

Исходя из полученных данных, было принято решение разработать информационную систему для научного сотрудника в виде веб-приложения.

#### 4.1.2. Анализ конкурентных технических решений

Проведем сравнительный анализ конкурентных технических решений, существующих на рынке информационных систем: «Здоровье», «Анализ данных». Анализируемые данные представлены в таблице 5.

Таблица 5 - Оценочная карта для сравнения конкурентных технических решений  
(разработок)

| Критерии оценки                                          | Вес критерия | Баллы          |                 |                 | Конкурентоспособность |                 |                 |
|----------------------------------------------------------|--------------|----------------|-----------------|-----------------|-----------------------|-----------------|-----------------|
|                                                          |              | Б <sub>ф</sub> | Б <sub>к1</sub> | Б <sub>к2</sub> | К <sub>ф</sub>        | К <sub>к1</sub> | К <sub>к2</sub> |
| 1                                                        | 2            | 3              | 4               | 5               | 6                     | 7               | 8               |
| <b>Технические критерии оценки ресурсоэффективности</b>  |              |                |                 |                 |                       |                 |                 |
| 1. Повышение производительности труда пользователя       | 0,13         | 5              | 4               | 5               | 0,65                  | 0,52            | 0,65            |
| 2. Удобство в эксплуатации                               | 0,05         | 5              | 5               | 4               | 0,25                  | 0,25            | 0,2             |
| 3. Надежность ИС                                         | 0,15         | 4              | 4               | 4               | 0,6                   | 0,6             | 0,6             |
| 4. Потребность в ресурсах памяти                         | 0,04         | 5              | 4               | 5               | 0,2                   | 0,16            | 0,2             |
| 5. Безопасность                                          | 0,07         | 5              | 5               | 5               | 0,35                  | 0,35            | 0,35            |
| 6. Функциональная мощность (предоставляемые возможности) | 0,08         | 4              | 3               | 4               | 0,32                  | 0,24            | 0,32            |
| 7. Простота эксплуатации                                 | 0,06         | 5              | 5               | 5               | 0,3                   | 0,3             | 0,3             |
| 8. Качество интеллектуального интерфейса                 | 0,1          | 5              | 4               | 4               | 0,5                   | 0,4             | 0,4             |
| <b>Экономические критерии оценки эффективности</b>       |              |                |                 |                 |                       |                 |                 |
| 1. Конкурентоспособность продукта                        | 0,1          | 5              | 3               | 4               | 0,5                   | 0,3             | 0,4             |
| 2. Цена                                                  | 0,15         | 5              | 2               | 3               | 0,75                  | 0,3             | 0,45            |
| 3. Срок выхода на рынок                                  | 0,01         | 3              | 5               | 5               | 0,03                  | 0,05            | 0,05            |
| 4. Предполагаемый срок эксплуатации                      | 0,01         | 5              | 3               | 4               | 0,05                  | 0,03            | 0,04            |
| 5. Послепродажное обслуживание                           | 0,01         | 5              | 4               | 3               | 0,05                  | 0,04            | 0,03            |
| 6. Финансирование научной разработки                     | 0,04         | 5              | 2               | 2               | 0,2                   | 0,08            | 0,08            |
| <b>Итого</b>                                             | <b>1</b>     | <b>66</b>      | <b>53</b>       | <b>57</b>       | <b>4,75</b>           | <b>3,62</b>     | <b>4,07</b>     |

Анализ конкурентных технических решений определяется по формуле 1:

$$K = \sum B_i F_i, \quad (1)$$

где  $K$  – конкурентоспособность научной разработки или конкурента;

$B_i$  – вес показателя (в долях единицы);

$F_i$  – балл  $i$ -го показателя.

Основываясь на знаниях о конкурентах, обратим внимание на сильные стороны разрабатываемой информационной системы:

- 1) Низкая цена продукта (в три раза ниже, чем у конкурентов).
- 2) Высокое качество интеллектуального интерфейса.
- 3) Длительный срок эксплуатации (5 лет гарантии, по сравнению с годом гарантии и конкурентов).

4) Качество послепродажного обслуживания (поддержка продукта в течение всего срока гарантии).

5) Для разработки продукта требовались минимальные вложения.

#### 4.1.3. Технология QuaD

Данная технология использована для анализа проекта, чтобы измерить характеристики качества этой разработки и ее перспективность использования в работе. В основе технологии QuaD лежит нахождение средневзвешенной величины двух групп показателей: качества и потенциала разработки [13]. Каждый показатель оценивается экспертным путем по сто балльной шкале. В таблице 6 представлена оценочная карта.

Таблица 6 – Оценочная карта для сравнения конкурентных технических решений

| Критерии оценки                                                       | Вес критерия | Баллы | Максимальный балл | Относительное значение | Средне взвешенное значение |
|-----------------------------------------------------------------------|--------------|-------|-------------------|------------------------|----------------------------|
| 1                                                                     | 2            | 3     | 4                 | 5                      | 6                          |
| <b>Показатели оценки качества разработки</b>                          |              |       |                   |                        |                            |
| Надежность                                                            | 0,2          | 70    | 100               | 0,7                    | 0,14                       |
| Функциональная мощность                                               | 0,2          | 90    | 100               | 0,9                    | 0,18                       |
| Простота эксплуатации                                                 | 0,3          | 90    | 100               | 0,9                    | 0,27                       |
| Качество интерфейса                                                   | 0,2          | 80    | 100               | 0,8                    | 0,16                       |
| <b>Показатели оценки коммерческого потенциала разработки</b>          |              |       |                   |                        |                            |
| Конкурентоспособность продукта по сравнению с существующими системами | 0,1          | 70    | 100               | 0,7                    | 0,7                        |
| <b>Итого:</b>                                                         | <b>1</b>     |       | <b>100</b>        |                        | <b>1,45</b>                |

Оценка качества и перспективности по технологии QuaD определяется по формуле

2:

$$P_{cp} = \sum B_i \cdot B_i \quad (2)$$

где  $P_{cp}$  – средневзвешенное значение показателя качества и перспективности научной разработки;

$B_i$  – вес показателя (в долях единицы);

$B_i$  – средневзвешенное значение  $i$ -го показателя.

Получаем, что  $P_{cp} = 1,45$ .

Значение  $P_{cp}$  позволяет говорить о перспективах разработки и качестве проведенного исследования. Учитывая, что  $P_{cp} = 1,45$ , можно утверждать, что данная разработка является перспективной.

#### 4.1.4. SWOT-анализ

SWOT-анализ является одним из самых распространенных методов анализа и стратегического планирования в менеджменте и маркетинге [14]. Он дает четкое представление факторах внешней и внутренней среды и указывает, в каких направлениях нужно действовать, используя сильные стороны, чтобы максимизировать возможности и свести к минимуму угрозы и слабые стороны. С помощью этого метода можно обозначить основные проблемы проекта, определить пути решения и перспективу развития.

Объектом для проведения SWOT-анализа является процесс использования информационной системы сбора и подготовки документов для проведения итоговой государственной аттестации выпускников.

Результатом анализа является разработка маркетинговой стратегии или гипотезы для дальнейшей проверки, они представлены в таблице 7.

Таблица 7 – SWOT-анализ

| 1                                                                                                                                                                                        | 2                                                                                                                                                                                                                                                                             | 3                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                                                                          | <b>Возможности:</b><br>1. Увеличение скорости проведения анализа медицинских данных<br>2. Исключение ошибок связанных с человеческим фактором.                                                                                                                                | <b>Угрозы:</b><br>1. Сбои в работе системы.<br>2. Сокращение времени будет недостаточным.<br>3. Изменение системы работы университета приведет за собой изменение работы системы.                                                                                                                                                                                                                                                |
| <b>Сильные стороны:</b><br>1. Упрощение работы с медицинскими данными.<br>2. Сокращение времени обработки медицинских данных.<br>3. Составление наиболее эффективной траектории лечения. | <b>Как воспользоваться возможностями</b><br>1. Хранить медицинские данные в электронном виде в необходимом для системы формате.<br>2. Быстрый и точный сбор статистики приведет к исключению работы людей с многочисленными данными и позволит получить информативные данные. | <b>За счет чего можно снизить угрозы</b><br>1. При разработке учитывать наиболее времязатратные этапы, сокращение времени должно быть не на проценты, а в разы сокращать время обработки.<br>2. Сбои в работе можно исключить за счет грамотного выбора разработчика и предоставления лучших инструментов для работы.<br>3. Проводить разработку системы таким образом, чтобы она была гибкой, легко восприимчивой к изменениям. |
| <b>Слабые стороны:</b><br>1. Восприимчивость системы к веб-серверам                                                                                                                      | <b>Что может помешать возможностям</b><br>1. Зависимость от сети                                                                                                                                                                                                              | <b>Самые большие опасности для фирмы</b><br>1. Разработанная системы не                                                                                                                                                                                                                                                                                                                                                          |

|                                  |                                                        |                                                                                                 |
|----------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| 2. Зависимость от сети Интернет. | интернет может привести к проблемам сохранения данных. | будет соответствовать требованиям.<br>2. Время, затраченное на ее создание, будет потеряно зря. |
|----------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------------------|

Таким образом, в результате SWOT-анализа были рассмотрены сильные и слабые стороны разработки информационной системы, выявлены возможные перспективы ее создания и рассмотрены варианты минимизации влияния угроз, которые могут этому помешать. Для повышения эффективности работы приложения необходимо минимизировать ошибки, контролировать процесс создания и проверять работу системы в соответствии с поставленными требованиями.

## 4.2. Планирование научно-исследовательских работ

### 4.2.1. Структура работ в рамках научного исследования

Планирование комплекса предполагаемых работ осуществляется в следующем порядке:

- определение структуры работ в рамках научного исследования;
- определение участников каждой работы;
- установление продолжительности работ;
- построение графика проведения научных исследований и разработок.

Для выполнения научного исследования сформирована рабочая группа, в состав которой входят студенты и преподаватель – руководитель проекта. По каждому виду запланированных работ установлена соответствующая должность исполнителей, эти данные представлены в таблице 8.

Таблица 8 - Перечень этапов, работ и распределение исполнителей

| Основные этапы                                                              | № раб. | Содержание работ                               | Должность исполнителя         |
|-----------------------------------------------------------------------------|--------|------------------------------------------------|-------------------------------|
| Оформление идеи разработки                                                  | 1      | Продумать идею и записать основные данные      | Руководитель проекта, студент |
| Разработка технического задания                                             | 2      | Подробное описание идеи                        | Руководитель, студент         |
|                                                                             | 3      | Выявление функциональных блоков                | Руководитель, студент         |
|                                                                             | 4      | Назначение ответственных за каждый блок        | Руководитель                  |
|                                                                             | 5      | Календарное планирование работ                 | Руководитель                  |
| Поиск подходов для решения текущих задач, написание кода и его тестирование | 6      | Поиск решения по каждому функциональному блоку | Студент                       |
|                                                                             | 7      | Реализация найденного решения                  | Студент                       |
|                                                                             | 8      | Тестирование                                   | Студент                       |

|                                                          |    |                                                                    |                       |
|----------------------------------------------------------|----|--------------------------------------------------------------------|-----------------------|
|                                                          | 9  | Отладка: устранение ошибок и повторное тестирование блоков проекта | Студент               |
| Обобщение и оценка результатов                           | 10 | Оценка эффективности полученных результатов                        | Руководитель, студент |
|                                                          | 11 | Определение целесообразности проведения ОКР                        | Руководитель, студент |
| Разработка технической документации и проектирование     | 12 | Разработка UML диаграмм по проекту                                 | Студент               |
|                                                          | 13 | Выбор проектируемой системы                                        | Руководитель, студент |
|                                                          | 14 | Оценка эффективности информационной системы                        | Руководитель, студент |
| Тестирование системы конечными пользователями            | 15 | Тестирование конечными пользователями                              | Студент               |
|                                                          | 16 | Исправление ошибок, доработка интерфейса                           | Студент               |
| Оформление отчета по НИР (комплекта документации по ОКР) | 17 | Составление пояснительной записки                                  | Руководитель, студент |

#### 4.2.2. Определение трудоемкости выполнения работ

Трудовые затраты в большинстве случаев образуют основную часть стоимости разработки, поэтому важным моментом является определение трудоемкости работ каждого из участников научного исследования.

Трудоемкость выполнения научного исследования оценивается экспертным путем в человеко-днях и носит вероятностный характер, т.к. зависит от множества трудно учитываемых факторов [15]. Для определения ожидаемого (среднего) значения трудоемкости  $t_{ожі}$  используется формула 3:

$$t_{ожі} = \frac{3t_{mini} + 2t_{maxi}}{5}, \quad (3)$$

где  $t_{ожі}$  – ожидаемая трудоемкость выполнения  $i$ -ой работы чел.-дн.;

$t_{mini}$  – минимально возможная трудоемкость выполнения заданной  $i$ -ой работы (оптимистическая оценка: в предположении наиболее благоприятного стечения обстоятельств), чел.-дн.;

$t_{maxi}$  – максимально возможная трудоемкость выполнения заданной  $i$ -ой работы (пессимистическая оценка: в предположении наиболее неблагоприятного стечения обстоятельств), чел.-дн.

Исходя из ожидаемой трудоемкости работ, определяется продолжительность каждой работы в рабочих днях  $T_p$ , учитывающая параллельность выполнения работ несколькими исполнителями. Такое вычисление необходимо для обоснованного расчета заработной

платы, так как удельный вес зарплаты в общей сметной стоимости научных исследований составляет около 65 %. Продолжительно одной работы определяется по формуле 4:

$$T_{pi} = \frac{t_{ожі}}{ч_i}, \quad (4)$$

где  $T_{pi}$  – продолжительность одной работы, раб. дн.;

$t_{ожі}$  – ожидаемая трудоемкость выполнения одной работы, чел.-дн.

$Ч_i$  – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Составим таблицу 9 и подсчитаем значения  $t_{ожі}$  и  $T_{pi}$  для каждой из работ.

Таблица 9 – Перечень работ и подсчет показателей

| Перечень                                                           | t min | t max | t ожид | Ч | Tpi  |
|--------------------------------------------------------------------|-------|-------|--------|---|------|
| Продумать идею и записать основные данные                          | 4     | 7     | 5,2    | 3 | 1,7  |
| Подробное описание идеи                                            | 6     | 8     | 6,8    | 3 | 2,3  |
| Выявление функциональных блоков                                    | 2     | 4     | 2,8    | 3 | 0,9  |
| Назначение ответственных за каждый блок                            | 1     | 5     | 2,6    | 1 | 2,6  |
| Календарное планирование работ                                     | 2     | 7     | 4      | 1 | 4    |
| Поиск решения по каждому функциональному блоку                     | 5     | 10    | 7      | 2 | 3,5  |
| Реализация решений                                                 | 60    | 90    | 72     | 2 | 36   |
| Тестирование                                                       | 10    | 14    | 11,6   | 2 | 5,8  |
| Отладка: устранение ошибок и повторное тестирование блоков проекта | 5     | 15    | 9      | 2 | 4,5  |
| Оценка эффективности полученных результатов                        | 6     | 12    | 8,4    | 3 | 2,8  |
| Определение целесообразности проведения ОКР                        | 3     | 7     | 4,6    | 3 | 1,5  |
| Разработка UML диаграмм по проекту                                 | 4     | 11    | 6,8    | 2 | 3,4  |
| Выбор проектируемой системы                                        | 1     | 3     | 1,8    | 3 | 0,6  |
| Оценка эффективности информационной системы                        | 3     | 6     | 4,2    | 3 | 1,4  |
| Тестирование конечными пользователями                              | 14    | 23    | 17,6   | 5 | 3,52 |
| Исправление ошибок, доработка интерфейса                           | 20    | 45    | 30     | 2 | 15   |
| Написание пояснительной записки                                    | 14    | 30    | 20,4   | 3 | 6,8  |

#### 4.2.3. Разработка графика проведения научного исследования

Составим график проведения научных работ в форме диаграммы Ганта.

**Диаграмма Ганта** – горизонтальный ленточный график, на котором работы по теме представляются протяженными во времени отрезками, характеризующимися датами начала и окончания выполнения данных работ [13].

Для удобства построения графика, длительность каждого из этапов работ из рабочих дней переведем в календарные дни. Для этого воспользуемся формулой 5:

$$T_{ki} = T_{pi} \cdot k_{\text{кал}}, \quad (5)$$

где  $T_{ki}$  – продолжительность выполнения  $i$ -й работы в календарных днях;

$T_{pi}$  – продолжительность выполнения  $i$ -й работы в рабочих днях;

$k_{\text{кал}}$  – коэффициент календарности.

Коэффициент календарности определяется по формуле 6:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}}, \quad (6)$$

где  $T_{\text{кал}}$  – количество календарных дней в году;

$T_{\text{вых}}$  – количество выходных дней в году;

$T_{\text{пр}}$  – количество праздничных дней в году.

$$k_{\text{кал}} = 365 / (365 - 125) = 1,5.$$

Рассчитанные значения в календарных днях по каждой работе  $T_{ki}$  округлим до целого числа.

Все рассчитанные значения сведены в таблицу 6. Исполнители: Р-руководитель, С-студент. Временные параметры рассчитаны для трех исполнений.

Исполнение 1 – Данный проект.

Исполнение 2 – «Здоровье».

Исполнение 3 – «Анализ данных».

Таблица 10 - Временные показатели проведения научного исследования

| Название работы, исполнители                      | Трудоёмкость работ  |       |       |                     |       |       |                           |       |       | Длительность работ в рабочих днях $T_{pi}$ |       |       | Длительность работ в календарных днях $T_{ki}$ |       |       |
|---------------------------------------------------|---------------------|-------|-------|---------------------|-------|-------|---------------------------|-------|-------|--------------------------------------------|-------|-------|------------------------------------------------|-------|-------|
|                                                   | $t_{min}$ , чел-дни |       |       | $t_{max}$ , чел-дни |       |       | $t_{ож\bar{i}}$ , чел-дни |       |       |                                            |       |       |                                                |       |       |
|                                                   | Исп.1               | Исп.2 | Исп.3 | Исп.1               | Исп.2 | Исп.3 | Исп.1                     | Исп.2 | Исп.3 | Исп.1                                      | Исп.2 | Исп.3 | Исп.1                                          | Исп.2 | Исп.3 |
| Продумать идею и записать основные данные. Р,С    | 4                   | 4     | 5     | 7                   | 7     | 8     | 5,2                       | 5,2   | 6,2   | 1,7                                        | 1,7   | 2,1   | 3                                              | 3     | 3     |
| Подробное описание идеи. Р,С                      | 6                   | 6     | 7     | 8                   | 8     | 9     | 6,8                       | 6,8   | 7,8   | 2,3                                        | 2,3   | 2,6   | 3                                              | 3     | 4     |
| Выявление функциональных блоков. Р,С              | 2                   | 3     | 3     | 4                   | 5     | 5     | 2,8                       | 3,8   | 3,8   | 0,9                                        | 1,3   | 1,3   | 1                                              | 2     | 2     |
| Назначение ответственных за каждый блок. Р        | 1                   | 1     | 1     | 5                   | 5     | 5     | 2,6                       | 2,6   | 2,6   | 2,6                                        | 2,6   | 2,6   | 4                                              | 4     | 4     |
| Календарное планирование работ. Р                 | 2                   | 2     | 2     | 7                   | 7     | 7     | 4                         | 4     | 4     | 4,0                                        | 4,0   | 4,0   | 6                                              | 6     | 6     |
| Поиск решения по каждому функциональному блоку. С | 5                   | 5     | 5     | 10                  | 10    | 10    | 7                         | 7     | 7     | 3,5                                        | 3,5   | 3,5   | 5                                              | 5     | 5     |
| Реализация найденного решения. С                  | 60                  | 65    | 62    | 90                  | 95    | 92    | 72                        | 77    | 74    | 36,0                                       | 38,5  | 37,0  | 54                                             | 58    | 56    |
| Тестирование. С                                   | 10                  | 10    | 10    | 14                  | 14    | 14    | 11,6                      | 11,6  | 11,6  | 5,8                                        | 5,8   | 5,8   | 9                                              | 9     | 9     |

Продолжение таблицы 10

|                                                                       |            |            |            |            |            |            |              |              |              |             |             |             |            |            |            |
|-----------------------------------------------------------------------|------------|------------|------------|------------|------------|------------|--------------|--------------|--------------|-------------|-------------|-------------|------------|------------|------------|
| Отладка: устранение ошибок и повторное тестирование блоков проекта. С | 5          | 4          | 6          | 15         | 14         | 16         | 9            | 8            | 10           | 4,5         | 4,0         | 5,0         | 7          | 6          | 8          |
| Оценка эффективности полученных результатов. Р,С                      | 6          | 6          | 6          | 12         | 12         | 12         | 8,4          | 8,4          | 8,4          | 2,8         | 2,8         | 2,8         | 4          | 4          | 4          |
| Определение целесообразности проведения ОКР. Р,С                      | 3          | 4          | 4          | 7          | 8          | 8          | 4,6          | 5,6          | 5,6          | 1,5         | 1,9         | 1,9         | 2          | 3          | 3          |
| Разработка UML диаграмм по проекту. С                                 | 4          | 4          | 4          | 11         | 11         | 11         | 6,8          | 6,8          | 6,8          | 3,4         | 3,4         | 3,4         | 5          | 5          | 5          |
| Выбор проектируемой системы. Р,С                                      | 1          | 1          | 1          | 3          | 3          | 3          | 1,8          | 1,8          | 1,8          | 0,6         | 0,6         | 0,6         | 1          | 1          | 1          |
| Оценка эффективности информационной системы. Р,С                      | 3          | 3          | 4          | 6          | 6          | 7          | 4,2          | 4,2          | 5,2          | 1,4         | 1,4         | 1,7         | 2          | 2          | 3          |
| Тестирование конечными пользователями. С                              | 14         | 15         | 14         | 23         | 24         | 23         | 17,6         | 18,6         | 17,6         | 3,5         | 3,7         | 3,5         | 5          | 6          | 5          |
| Исправление ошибок, доработка интерфейса. С                           | 20         | 20         | 22         | 45         | 45         | 47         | 30           | 30           | 32           | 15,0        | 15,0        | 16,0        | 23         | 23         | 24         |
| Составление пояснительной записки Р,С                                 | 6          | 6          | 6          | 12         | 12         | 12         | 8,4          | 8,4          | 8,4          | 2,8         | 2,8         | 2,8         | 4          | 4          | 4          |
| <b>Итого:</b>                                                         | <b>152</b> | <b>159</b> | <b>162</b> | <b>279</b> | <b>286</b> | <b>289</b> | <b>202,8</b> | <b>209,8</b> | <b>212,8</b> | <b>92,4</b> | <b>95,3</b> | <b>96,6</b> | <b>139</b> | <b>143</b> | <b>145</b> |

По данным расчетам, программа будет разработана:

- в первом исполнении 139 дней
- во втором исполнении 143 дней
- в третьем исполнении 145 дней

Следовательно, можно сделать вывод, что в первом исполнении работы будет выполнена быстрее.

На основе таблицы 6 построен календарный план-график, представленный в таблице 11. График строился для максимального по длительности исполнения работ в рамках научно-исследовательского проекта с разбивкой по месяцам и декадам (10 дней) за период времени дипломирования. При этом работы на графике выделены различной штриховкой в зависимости от исполнителей, ответственных за ту или иную работу.

Таблица 11 - Календарный план-график проведения НИОКР по теме

| №      | Вид работ                                                             | $T_{ki}$ ,<br>кал. дн. | Продолжительность выполнения работ                                                                                                                                                                   |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
|--------|-----------------------------------------------------------------------|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|        |                                                                       |                        | февр.                                                                                                                                                                                                |                                                                                     |                                                                                     | март                                                                                |                                                                                     |                                                                                     | апрель                                                                              |                                                                                     |                                                                                     | май                                                                                 |                                                                                       |                                                                                       | июнь                                                                                  |                                                                                       |
|        |                                                                       |                        | 1                                                                                                                                                                                                    | 2                                                                                   | 3                                                                                   | 1                                                                                   | 2                                                                                   | 3                                                                                   | 1                                                                                   | 2                                                                                   | 3                                                                                   | 1                                                                                   | 2                                                                                     | 3                                                                                     | 1                                                                                     | 2                                                                                     |
| 1      | Продумать идею и записать основные данные. Р,С                        | 3                      |                                                                                                                   |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 2      | Подробное описание идеи. Р,С                                          | 3                      |                                                                                                                   |  |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 3      | Выявление функциональных блоков. Р,С                                  | 1                      |                                                                                                                   |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 4      | Назначение ответственных за каждый блок. Р                            | 4                      |                                                                                                                                                                                                      |  |  |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 5      | Календарное планирование работ. Р                                     | 6                      |                                                                                                                                                                                                      |                                                                                     |  |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 6      | Поиск решения по каждому функциональному блоку. С                     | 5                      |                                                                                                                                                                                                      |                                                                                     |  |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 7      | Реализация найденного решения. С                                      | 54                     |                                                                                                                                                                                                      |                                                                                     |  |  |  |  |  |  |  |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 8      | Тестирование. С                                                       | 9                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |  |  |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 9      | Отладка: устранение ошибок и повторное тестирование блоков проекта. С | 7                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |  |  |                                                                                       |                                                                                       |                                                                                       |                                                                                       |
| 10     | Оценка эффективности полученных результатов. Р,С                      | 4                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |  |    |                                                                                       |                                                                                       |                                                                                       |
| 11     | Определение целесообразности проведения ОКР. Р,С                      | 2                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |    |    |                                                                                       |                                                                                       |
| 12     | Разработка UML диаграмм по проекту. С                                 | 5                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |  |  |                                                                                       |                                                                                       |
| 13     | Выбор проектируемой системы. Р,С                                      | 1                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |  |                                                                                       |                                                                                       |
| 14     | Оценка эффективности информационной системы. Р,С                      | 2                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |  |  |                                                                                       |
| 15     | Тестирование конечными пользователями. С                              | 5                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |  |  |                                                                                       |
| 16     | Исправление ошибок, доработка интерфейса. С                           | 23                     |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |  |  |  |
| 17     | Составление пояснительной записки Р,С                                 | 4                      |                                                                                                                                                                                                      |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |  |
| Итого: |                                                                       |                        |  – Студент  – Руководитель |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                     |                                                                                       |                                                                                       |                                                                                       |                                                                                       |

### 4.3. Бюджет научно-технического исследования (НТИ)

При планировании бюджета НТИ должно быть обеспечено полное и достоверное отражение всех видов расходов, связанных с его выполнением [14]. В процессе формирования бюджета НТИ используется следующая группировка затрат по статьям:

- материальные затраты НТИ;
- расчет амортизации;
- основная заработная плата исполнителей темы;
- дополнительная заработная плата исполнителей темы;
- отчисления во внебюджетные фонды (страховые отчисления);
- накладные расходы.

#### 4.3.1. Расчет материальных затрат НТИ

Расчет материальных затрат осуществляется по формуле 6:

$$З_m = (1 + k_T) \cdot \sum_{i=1}^m C_i \cdot N_{расхi}, \quad (7)$$

где  $m$  – количество видов материальных ресурсов, потребляемых при выполнении научного исследования;

$N_{расхi}$  – количество материальных ресурсов  $i$ -го вида, планируемых к использованию при выполнении научного исследования (шт., кг, м, м<sup>2</sup> и т.д.);

$C_i$  – цена приобретения единицы  $i$ -го вида потребляемых материальных ресурсов (руб./шт., руб./кг, руб./м, руб./м<sup>2</sup> и т.д.);

$k_T$  – коэффициент, учитывающий транспортно-заготовительные расходы.

Для разработки данного продукта необходимы следующие материальные ресурсы:

- Системный блок
- Принтер
- Монитор

Расчет материальных затрат представлен в таблице 12.

Таблица 12 - Материальные затраты

| Наименование   | Ед.<br>изм. | Количество |           |           | Цена за ед., |       |       | Затраты на материалы,<br>(З <sub>м</sub> ), руб. |       |       |
|----------------|-------------|------------|-----------|-----------|--------------|-------|-------|--------------------------------------------------|-------|-------|
|                |             |            |           |           | руб.         |       |       |                                                  |       |       |
|                |             | Ис<br>п.1  | Ис<br>п.2 | Ис<br>п.3 | Исп.1        | Исп.2 | Исп.3 | Исп.1                                            | Исп.2 | Исп.3 |
| Системный блок | шт.         | 1          | 1         | 1         | 34290        | 35800 | 34900 | 34290                                            | 35800 | 34900 |
| Принтер        |             | 1          | 1         | 1         | 8550         | 8000  | 7520  | 8550                                             | 8000  | 7520  |
| Монитор        |             | 2          | 2         | 2         | 7 690        | 8200  | 8500  | 15380                                            | 16400 | 17000 |
| Итого:         |             |            |           |           |              |       |       | 58220                                            | 60200 | 59420 |

#### 4.3.2. Основная заработная плата исполнителей темы

Основная заработная плата рассчитывается по формуле 8:

$$C_{\text{осн/зп}} = \sum_{i=1}^n t_i \cdot C_{\text{зп}i} \quad (8)$$

где  $n$  – количество видов работ;

$t_i$  – затраты труда на выполнение  $i$ -го вида работ, в днях;

$C_{\text{зп}i}$  – среднедневная заработная плата работника, выполняющего  $i$ -ый вид работы, руб./день.

Среднедневная заработная плата рассчитывается по формуле 9:

$$C_{\text{зп}i} = \frac{D \cdot K \cdot M_p}{F_o} \quad (9)$$

где  $D$  – месячный должностной оклад работника;

$K$  – коэффициент, учитывающий коэффициент по премиям и районный коэффициент ( $K=1,3$ );

$M_p$  – количество месяцев работы без отпуска в течение года;

$F_o$  – действительный годовой фонд рабочего времени работника, в днях.

При отпуске 28 дня  $M_p=11,08$ .

Результаты расчета действительного годового фонда проведены в таблице 12.

Таблица 12 - Годовой фонд рабочего времени

| Показатели рабочего времени, дни       | Исполнение 1 | Исполнение 2 | Исполнение 3 |
|----------------------------------------|--------------|--------------|--------------|
| Календарное число дней в году          | 365          | 365          | 365          |
| Количество нерабочих дней              |              |              |              |
| Выходные                               | 104          | 104          | 104          |
| Праздники (фактически по каждому году) | 13           | 13           | 13           |
| Планируемые потери отпуска             | 28           | 28           | 28           |
| Действительный годовой фонд            | 220          | 220          | 220          |

Расчет затрат на основную заработную плату приведен в таблице 13. При этом затраты на оплату труда студента-дипломника определяются как оклад студента ( $D = 13952,44$ ), а оклад руководителя проекта составляет 23264,86. Коэффициент  $K$ , учитывающий районный коэффициент равный 1,3.

Таблица 13 - Затраты на основную заработную плату

| Исполнители  | Среднедневная заработная плата $C_{\text{зп}}$ (руб.) |       |       | Трудоемкость ( $t_i$ ), чел-дни |       |       | Затраты на основную зарплату (руб.) |                 |                 |
|--------------|-------------------------------------------------------|-------|-------|---------------------------------|-------|-------|-------------------------------------|-----------------|-----------------|
|              | Исп 1                                                 | Исп 2 | Исп 3 | Исп 1                           | Исп 2 | Исп 3 | Исп 1                               | Исп 2           | Исп 3           |
| Руководитель | 1523,2                                                |       |       | 20,6                            | 21,4  | 22,4  | 31377,92                            | 32596,48        | 34119,68        |
| Студент      | 913,4                                                 |       |       | 85,7                            | 88,7  | 90    | 78278,38                            | 81018,58        | 82206           |
| <b>Итого</b> |                                                       |       |       |                                 |       |       | <b>109656,3</b>                     | <b>113615,1</b> | <b>116325,7</b> |

#### 4.3.3. Дополнительная заработная плата

Дополнительная заработная плата включает заработную плату за не отработанное рабочее время, но гарантированную действующим законодательством [15].

Расчет дополнительной заработной платы ведется по формуле 10:

$$З_{доп} = k_{доп} \cdot З_{осн} \quad (10)$$

где  $k_{доп}$  – коэффициент дополнительной заработной платы (на стадии проектирования принимается равным 0,12 – 0,15).

$k_{доп}$  равен 0,12. Результаты по расчетам дополнительной заработной платы сведены в таблицу 14.

Таблица 14 - Затраты на дополнительную заработную плату

| Исполнители  | Основная<br>зарплата( руб.) |          |          | Коэффициент<br>дополнительной<br>заработной<br>платы ( $k_{доп}$ ) | Дополнительная<br>зарплата( руб.) |                 |                 |
|--------------|-----------------------------|----------|----------|--------------------------------------------------------------------|-----------------------------------|-----------------|-----------------|
|              | Исп.1                       | Исп.2    | Исп.3    |                                                                    | Исп.1                             | Исп.2           | Исп.3           |
| Руководитель | 31377,92                    | 32596,48 | 34119,68 | 0,12                                                               | 3765,35                           | 3911,57         | 4094,36         |
| Студент      | 78278,38                    | 81018,58 | 82206    | 0,12                                                               | 9393,4                            | 9722,24         | 9864,72         |
| <b>Итого</b> |                             |          |          |                                                                    | <b>13158,76</b>                   | <b>13633,81</b> | <b>13959,08</b> |

#### 4.3.4. Отчисления во внебюджетные фонды (страховые отчисления)

Величина отчислений во внебюджетные фонды определяется исходя из формулы 11:

$$З_{внеб} = k_{внеб} \cdot (З_{осн} + З_{доп}), \quad (11)$$

где  $k_{внеб}$  – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

На основании пункта 1 ст.58 закона №212-ФЗ для учреждений осуществляющих образовательную и научную деятельность в 2014 году вводится пониженная ставка страховых взносов — 27,1%.

Отчисления во внебюджетные фонды представлены в таблице 15.

Таблица 15- Отчисления во внебюджетные фонды

| Исполнитель                                  | Основная заработная плата, руб. |          |          | Дополнительная заработная плата, руб. |         |         |
|----------------------------------------------|---------------------------------|----------|----------|---------------------------------------|---------|---------|
|                                              | Исп. 1                          | Исп. 2   | Исп. 3   | Исп. 1                                | Исп. 2  | Исп. 3  |
| Руководитель проекта                         | 31377,92                        | 32596,48 | 34119,68 | 3765,35                               | 3911,57 | 4094,36 |
| Студент                                      | 78278,38                        | 81018,58 | 82206    | 9393,4                                | 9722,24 | 9864,72 |
| Коэффициент отчислений во внебюджетные фонды | 27,1%                           |          |          |                                       |         |         |
| Итого                                        |                                 |          |          |                                       |         |         |
| Исполнение 1                                 | 36844,52                        |          |          |                                       |         |         |
| Исполнение 2                                 | 38174,66                        |          |          |                                       |         |         |
| Исполнение 3                                 | 39085,43                        |          |          |                                       |         |         |

#### 4.3.5. Накладные расходы

Накладные расходы учитывают прочие затраты организации, не попавшие в предыдущие статьи расходов: печать и ксерокопирование материалов исследования, оплата услуг связи, электроэнергии, почтовые и телеграфные расходы, размножение материалов и т.д. Их величина определяется по формуле 12:

$$Z_{\text{накл}} = (\text{сумма статей } 1 \div 7) \cdot k_{\text{нр}}, \quad (12)$$

где  $k_{\text{нр}}$  – коэффициент, учитывающий накладные расходы.

Величина коэффициента накладных расходов равна 50%.

**Исполнение 1:**  $(58220/7) \cdot 0,5 = 4158,6$

**Исполнение 2:**  $(60200/7) \cdot 0,5 = 4300$

**Исполнение 3:**  $(59420/7) \cdot 0,5 = 4244,3$

#### 4.3.6. Формирование бюджета затрат научно-исследовательского проекта

Рассчитанная величина затрат научно-исследовательской работы (темы) является основой для формирования бюджета затрат проекта, который при формировании договора с заказчиком защищается научной организацией в качестве нижнего предела затрат на разработку научно-технической продукции [15].

Определение бюджета затрат на научно-исследовательский проект по каждому варианту исполнения приведен в таблице 16.

Таблица 16 - Расчет бюджета затрат НТИ

| Наименование статьи                                             | Сумма, руб.     |                 |                 |
|-----------------------------------------------------------------|-----------------|-----------------|-----------------|
|                                                                 | Исп.1           | Исп.2           | Исп.3           |
| 1. Материальные затраты НТИ                                     | 58220           | 60200           | 59420           |
| 2. Затраты по основной заработной плате исполнителей темы       | 109656,3        | 113615,1        | 116325,7        |
| 3. Затраты по дополнительной заработной плате исполнителей темы | 13158,76        | 13633,81        | 13959,08        |
| 4. Отчисления во внебюджетные фонды                             | 36844,52        | 38174,66        | 39085,43        |
| 5. Накладные расходы                                            | 4158,6          | 4300            | 4244,3          |
| 6. Бюджет затрат НТИ                                            | <b>222038,2</b> | <b>229923,6</b> | <b>233034,5</b> |

**Вывод:** Основываясь на данных, полученных в пунктах 4.3.1 – 4.3.5, был рассчитан бюджет затрат научно-исследовательской работы для трех исполнителей. Наиболее низким по себестоимости оказался проект первого исполнителя, затраты на его полную реализацию составляют **222038,2** рублей.

#### 4.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Определение эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Его нахождение связано с определением двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования получают в ходе оценки бюджета затрат трех (или более) вариантов исполнения научного исследования.

**Интегральный финансовый показатель** разработки определяется по формуле 13:

$$I_{\text{финр}}^{\text{исп.}i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}, \quad (13)$$

где  $I_{\text{финр}}^{\text{исп.}i}$  – интегральный финансовый показатель разработки;

$\Phi_{pi}$  – стоимость  $i$ -го варианта исполнения;

$\Phi_{\text{max}}$  – максимальная стоимость исполнения научно-исследовательского проекта (в т.ч. аналоги).

Расчет:

$$I_{\text{финр}}^{\text{исп.1}} = \frac{222038,2}{233034,5} = 0,95$$

$$I_{\text{финр}}^{\text{исп.2}} = \frac{229923,6}{233034,5} = 0,99$$

$$I_{\text{финр}}^{\text{исп.3}} = \frac{233034,5}{233034,5} = 1$$

Полученная величина интегрального финансового показателя разработки отражает соответствующее численное увеличение бюджета затрат разработки в разах (значение больше единицы), либо соответствующее численное удешевление стоимости разработки в разах (значение меньше единицы, но больше нуля).

Как видно из результатов, полученных выше, Исполнение 1 имеет наименьший интегральный финансовый показатель, а это значит, что данное исследование обладает наименьшей стоимостью разработки.

**Интегральный показатель ресурсоэффективности** вариантов исполнения объекта исследования можно определить по формуле 14:

$$I_{pi} = \sum a_i \cdot b_i, \quad (14)$$

где  $I_{pi}$  – интегральный показатель ресурсоэффективности для  $i$ -го варианта исполнения разработки;

$a_i$  – весовой коэффициент  $i$ -го варианта исполнения разработки;

$b_i$  – бальная оценка  $i$ -го варианта исполнения разработки, устанавливается экспертным путем по выбранной шкале оценивания;

$n$  – число параметров сравнения.

Расчет интегрального показателя ресурсоэффективности приведен в таблице 17.

Таблица 17 - Сравнительная оценка характеристик вариантов исполнения проекта

| <b>Объект исследования</b><br><b>Критерии</b>                       | <b>Весовой коэффициент параметра</b> | <b>Исп.1</b> | <b>Исп.2</b> | <b>Исп.3</b> |
|---------------------------------------------------------------------|--------------------------------------|--------------|--------------|--------------|
| 1. Способствует росту производительности труда пользователя         | 0,2                                  | 5            | 3            | 4            |
| 2. Удобство в эксплуатации (соответствует требованиям потребителей) | 0,15                                 | 4            | 2            | 3            |
| 3. Помехоустойчивость                                               | 0,15                                 | 5            | 3            | 3            |
| 4. Энергосбережение                                                 | 0,1                                  | 4            | 3            | 3            |
| 5. Надежность                                                       | 0,25                                 | 4            | 4            | 4            |
| 6. Материалоемкость                                                 | 0,15                                 | 4            | 4            | 4            |
| <b>ИТОГО:</b>                                                       | <b>1</b>                             | <b>4,04</b>  | <b>3,15</b>  | <b>3,6</b>   |

$$I_{p-исп1} = 5 \cdot 0,2 + 4 \cdot 0,15 + 5 \cdot 0,15 + 4 \cdot 0,1 + 4 \cdot 0,25 + 5 \cdot 0,05 + 4 \cdot 0,01 = 4,04;$$

$$I_{p-исп2} = 3 \cdot 0,2 + 2 \cdot 0,15 + 3 \cdot 0,15 + 3 \cdot 0,1 + 4 \cdot 0,25 + 2 \cdot 0,05 + 4 \cdot 0,1 = 3,15;$$

$$I_{p-исп3} = 4 \cdot 0,2 + 3 \cdot 0,15 + 3 \cdot 0,15 + 3 \cdot 0,1 + 4 \cdot 0,25 + 4 \cdot 0,05 + 4 \cdot 0,1 = 3,6.$$

Исходя из вычислений, показатель ресурсоэффективности Исполнения 1 имеет

достаточно высокое значение (по 5-бальной шкале), что говорит об эффективности использования технического проекта. Высокие баллы надежности и помехоустойчивости позволяют судить о надежности системы. НТИ-2 и НТИ-3 уступают НТИ-1 в данном показателе, что свидетельствует о низкой ресурсоэффективности данных исследований в сравнении с Исполнением 1.

**Интегральный показатель эффективности вариантов исполнения разработки ( $I_{исп.i}$ )**

$$I_{исп.1} = \frac{I_{р-исп1}}{I_{финр}} = \frac{4.04}{0.95} = 4.3;$$

$$I_{исп.2} = \frac{I_{р-исп2}}{I_{финр}} = \frac{3.15}{0.99} = 3.2;$$

$$I_{исп.3} = \frac{I_{р-исп3}}{I_{финр}} = \frac{3.6}{1} = 3.6.$$

Исходя из полученных данных, можно сделать вывод, что эффективность исполнения у НТИ-1 выше, чем у других (равно 4,3), на втором месте НТИ-3 (3,6), и самый низкий показатель у НТИ-2 (3,2).

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволяет определить сравнительную эффективность проекта и выбрать наиболее целесообразный вариант из предложенных. Сравнительная эффективность проекта ( $\mathcal{E}_{cp}$ ) рассчитывается по формуле 15:

$$\mathcal{E}_{cp} = \frac{I_{исп.1}}{I_{исп.2}} \quad (15)$$

Сравнительная эффективность разработки, представлена в таблице 18.

Таблица 18 - Сравнительная эффективность разработки

| № п/п | Показатели                                              | Исп.1 | Исп.2 | Исп.3 |
|-------|---------------------------------------------------------|-------|-------|-------|
| 1     | Интегральный финансовый показатель разработки           | 0,95  | 0,99  | 1     |
| 2     | Интегральный показатель ресурсоэффективности разработки | 4,04  | 3,15  | 3,6   |
| 3     | Интегральный показатель эффективности                   | 4,3   | 3,2   | 3,6   |
| 4     | Сравнительная эффективность вариантов исполнения        | 1     | 0,7   | 0,8   |

### Выводы по разделу:

Таким образом, сравнительная эффективность у НТИ-1 равно 1, что говорит о самой высокой эффективности среди предложенных вариантов исследований. НТИ-2 отстает по данному показателю от НТИ-1 на 0,3, а НТИ-3 на 0,2.

В ходе работы был определен целевой рынок проекта, построена карта

сегментирования, выявлено какие ниши на рынке услуг по разработке интернет-ресурсов не заняты конкурентами или где уровень конкуренции низок.

Произведена оценка качества и перспективности по технологии *QuaD*, где было выявлено, что данная разработка является перспективной.

Произведен *SWOT*-анализа, в ходе которого были рассмотрены сильные и слабые стороны разработки информационной системы, выявлены возможные перспективы ее создания и рассмотрены варианты минимизации влияния угроз, которые могут этому помешать. Для повышения эффективности работы приложения необходимо минимизировать ошибки, контролировать процесс создания и проверять работу системы в соответствии с поставленными требованиями.

В рамках анализа конкурентных технических решений, определена трудоемкость данного проекта, она равна 139 дням. Это на 4 дня меньше, чем у НТИ-2, и на 6 дней меньше, чем у НТИ-3, что говорит о том, что данный проект будет выполнен быстрее, чем его конкурентные варианты. На основании трудоемкости была построена диаграмма Ганта.

Так же были рассчитаны величины затрат научно-исследовательских работ. В результате проведенных расчетов, бюджет затрат НТИ Исполнения 1 составил **222038,2**. Бюджет НТИ Исполнения 2 равен **229923,6**. 3-его Исполнения: **233034,5**. Отсюда следует сделать вывод о том, что первый исполнитель наименее затратный при реализации проекта. Первое исполнение лидирует и по показателю сравнительной эффективности, оно равно 1, в то время как для НТИ-2 этот показатель равен 0,7, а для НТИ-3 — 0,8. С позиций технической и финансовой ресурсоэффективности мы можем сделать выводы о том, что научно – техническое решение, представленное первым исполнителем, является более предпочтительным.



## **Глава 5 Социальная ответственность**

Направление данной ВРК работы - разработка информационной системы для анализа медицинских данных. Работа с системой заключается во взаимодействии научных работников с компьютером. Рабочая зона для выполнения ВКР – кабинет №204 Института Кибернетики Национального Исследовательского Томского Политехнического Университета. Рабочим местом научных работников являются оборудованный компьютерный стол и офисный стул.

Длительная работа за компьютером пагубно воздействует на здоровье человека. Компьютерное рабочее место и, прежде всего монитор, является источником электростатического поля; слабых электромагнитных излучений в низкочастотном и высокочастотном диапазонах (2 Гц...400 кГц); рентгеновского излучения; ультрафиолетового излучения; инфракрасного излучения; излучения видимого диапазона [16].

Человек, работающий за персональным компьютером, обычно сидит в неподвижной позе и напряжен. Это приводит к усталости и появлению болей в шее, плечевых суставах, позвоночнике. Наиболее сильная нагрузка приходится на зрительный аппарат человека. Длительная работа за клавиатурой может вызывать неприятные болевые ощущения в запястьях, локтевых суставах, кистях и пальцах рук.

Разработка данной методики не оказывает отрицательного влияния на общество и окружающую среду.

Данный раздел предназначен для разработки комплекса мер технического, организационного, режимного и правового характера, минимизирующих негативные последствия разработки информационной системы.

### **5.1. Техногенная безопасность**

В рамках разработки данной информационной системы были выявлены следующие факторы, отрицательно влияющие на человека:

- повышенная температура ПК;
- повышенный уровень электромагнитных излучений;
- недостаточность естественного освещения;
- недостаточность искусственного освещения;
- зрительное напряжение;
- повышенный уровень шума.

Высокая температура компьютера может привести к ожогам различной степени. Наиболее распространённая и менее опасная степень – первая. Ожоги данной степени в

большинстве случаев проходят через 2-4 дня. Однако, при нарушении правил пожарной безопасности, а также повреждений проводки возникает вероятность получить ожог третьей или четвертой степени.

Непостоянная температура в помещении может вызвать физический дискомфорт у сотрудников, что негативно повлияет на их производительности труда.

Санитарные правила и нормы 2.2.4.548-96 устанавливают определённый микроклимат для всех типов рабочих помещений. Они выполняют обязательную функцию для любой из возможных форм собственности. СанПиН 2.2.4.548-96 обязывает следовать безусловным правилам и требованиям, выполнение которых напрямую зависит от работодателя [16]. Категории работ и нормы указаны ниже в таблице 18. Допустимые нормы представлены в таблице 19.

Таблица 18- Оптимальные величины показателей микроклимата на рабочих местах производственных помещений

| Период года | Категория работ по уровню энергозатрат, Вт | Температура воздуха, °С | Температура поверхностей, °С | Относительная влажность воздуха, % | Скорость движения воздуха, м/с |
|-------------|--------------------------------------------|-------------------------|------------------------------|------------------------------------|--------------------------------|
| Холодный    | Ia (до 139)                                | 22 - 24                 | 21 - 25                      | 60 - 40                            | 0,1                            |
| Теплый      | Ia (до 139)                                | 23 - 25                 | 22 - 26                      | 60 - 40                            | 0,1                            |

Таблица 19 - Допустимые величины показателей микроклимата

| Период года | Категория работ по уровню энергозатрат, Вт | Температура воздуха, °С           |                                   | Температура поверхностей, °С | Относительная влажность воздуха, % | Скорость движения воздуха, м/с                                      |                                                                       |
|-------------|--------------------------------------------|-----------------------------------|-----------------------------------|------------------------------|------------------------------------|---------------------------------------------------------------------|-----------------------------------------------------------------------|
|             |                                            | диапазон ниже оптимальных величин | диапазон выше оптимальных величин |                              |                                    | для диапазона температур воздуха ниже оптимальных величин, не более | для диапазона температур воздуха выше оптимальных величин, не более** |
| Холодный    | Ia (до 139)                                | 20,0-21,9                         | 24,1-25,0                         | 19,0-26,0                    | 15-75*                             | 0,1                                                                 | 0,1                                                                   |
| Теплый      | Ia (до 139)                                | 21,0-22,9                         | 25,1-28,0                         | 20,0-29,0                    | 15-75*                             | 0,1                                                                 | 0,2                                                                   |

Температура воздуха в помещении в диапазоне от 23 до 26 ° С. Влажность воздуха – около 40 %. Скорость движения воздуха – 0,1 м/сек. Достигается это за счет использования средств местного кондиционирования воздуха, кроме этого помещение проветривается во время обеденного перерыва. Охлаждение сотрудников не происходит, так как помещение хорошо отапливается.

Параметры микроклимата соответствуют нормам, следовательно, в помещении соблюдаются необходимые для работы условия.

Среднесменная температура воздуха для разнообразных категорий работ не может превышать или быть ниже допустимой. В зависимости от того, насколько температура выше или ниже установленной, время нахождения на рабочем месте ограничивается до определённых часов [17].

Согласно статье 212 ТК РФ работодатель несёт ответственность за установление правильного микроклимата и непосредственно за здоровье своих сотрудников, а также он обязан обеспечить соответствующие нормы на каждом рабочем месте. Невыполнение требований ведёт к ответственности и закрепляется как нарушение трудового законодательства и ущемление прав работающих [18].

При работе за компьютером, пользователь получает вред от электромагнитных излучений. Электромагнитные излучения негативно влияют на сердечно-сосудистую систему, иммунную систему, на глаза. Беременным женщинам рекомендуется избегать большого влияния электромагнитных излучений. Электромагнитные излучения повреждают защитное биополе человека и отрицательно сказываются на его здоровье, повышая риск возникновения различного рода заболеваний. Напряженность электрического поля и плотность магнитного потока представлены в таблицах 20, 21 соответственно [19, 20].

Таблица 20 – Напряженность электрического поля

| Диапазон частот     | Напряженность электрического поля |
|---------------------|-----------------------------------|
| 5 Гц — 2 кГц, E1    | 25 В/м                            |
| 2 кГц — 400 кГц, E2 | 2,5 В/м                           |

Таблица 21 – Плотность магнитного потока

| Диапазон частот     | Плотность магнитного потока |
|---------------------|-----------------------------|
| 5 Гц — 2 кГц, B1    | 250 нТл                     |
| 2 кГц — 400 кГц, B2 | 25 нТл                      |

Наибольшее излучение от компьютера человек может получить не от монитора, а со стороны задней стенки. Поэтому следует ставить монитор к стене, тем самым не нанося вред здоровью коллег.

Для того, чтобы свести электромагнитное излучение к минимуму, системные блоки расположили настолько дальше от сотрудника, насколько это возможно, так же применяются жидкокристаллические мониторы, излучение которых намного ниже.

Мониторы расположены так, чтобы излучения поглощались не людьми, то есть в углу или у стены.

Согласно СанПиНу 2.2.1/2.1.1.1278-03 помещения с постоянным пребыванием людей должны иметь естественное освещение. Недостаточность естественного освещения так же отрицательно сказывается на производительности труда сотрудника, поскольку нехватка естественного света способствует к понижению гормона дофамина в мозгу человека [22].

Естественное освещение помещений важно, но зачастую его недостаточно для создания комфортных условий работы, труда и отдыха.

Поэтому его необходимо дополнять или заменять искусственным освещением. Помещение, освещенное посредством естественного и искусственного света одновременно, называется смешанным. Нормы искусственного и естественного освещения указаны в документе СНиП 23-05-95 [23]. Освещенность на поверхности стола в зоне размещения рабочего документа равна 300 - 500 лк. Освещенность поверхности экрана равна примерно 300 лк. Освещение не создает бликов на поверхности экрана. Яркость светильников общего освещения в зоне углов излучения от 50 до 90 градусов с вертикалью в продольной и поперечной плоскостях составляет не более 200 кд/м<sup>2</sup>. Все описанные показатели соответствуют нормам СанПиН 2.2.2/2.4.1340-03. КОЕ (коэффициент естественного освещения) – нормируемая величина, которая представляет собой отношение освещенности на рабочем месте к освещенности наружной:  $e=100E_p/E_n$ . Значение КЕО зависит от вида освещения и точности работ. Для работы в офисе (учебном кабинете) норма КЕО = 1,5, норма Кп = 10-15%.

Зрительное напряжение. Работа за компьютером причиняет вред глазам человека. Вред может быть зрительным: ухудшается качество зрения, дальность. Так же вред может быть связан с некомфортным ощущением в глазах. Недомогания начинаются после четырёх часов работы за компьютером. Наименьшую нагрузку на глаза дает считывание информации с монитора, чуть больше – ввод информации, а самая сильная утомляемость возникает при работе в диалоговом режиме и занятия (просмотр) компьютерной графикой.

Шум в окружающей среде создается одиночными или комплексными источниками, находящимися снаружи или внутри здания. Шум — это совокупность звуков различной интенсивности и частоты (шелест,дребезжание, скрип, визг и т. п.) [24]. Шум может создаваться оборудованием, людьми, городским транспортом. Программист работает в помещениях с низким уровнем общего шума. Источниками шумовых помех могут быть:

компьютер и периферийное оборудование к нему, вентиляционные установки и кондиционеры. Предельно допустимый уровень звука представлен в таблице 22.

Таблица 22 – Предельно допустимый уровень звука по СН 2.2.4/2.1.8.562-96

| Рабочие места                                   | Уровень звука, дБА |
|-------------------------------------------------|--------------------|
| Конструкторские бюро, программисты, лаборатории | 50                 |
| Помещения управления, рабочие комнаты           | 60                 |

Основным источником шума в кабинете и на рабочем месте являются вентиляторы блоков питания ЭВМ. Уровень шума колеблется от 35 до 40дБА.

Шумовая обстановка на рабочем месте соответствует норме. При регулярных проверках оборудования можно избежать превышения допустимого уровня шума.

На предприятии рабочее место сотрудника должно соответствовать определённым требованиям: Помещения с компьютерами необходимо оборудовать системами отопления, кондиционирования воздуха или эффективной приточно-вытяжной вентиляцией.

Для внутренней отделки интерьера помещений с компьютерами должны использоваться диффузно-отражающие материалы с коэффициентом отражения для потолка — 0,7-0,8; для стен — 0,5-0,6; для пола — 0,3-0,5 [25].

В помещениях эксплуатации компьютеров поверхность пола должна быть ровной, без выбоин, нескользкой, удобной для очистки и влажной уборки, обладать антистатическими свойствами.

В помещении должны находиться аптечка первой медицинской помощи, углекислотный огнетушитель.

К опасным факторам различного рода можно отнести следующее:

- механические опасности;
- термические опасности;
- электробезопасность;
- пожаровзрывобезопасность.

Рассмотрим более подробно каждый из описанных выше факторов.

Механические опасности, возникающие при работе с ПК в офисе, могут вызвать тяжелые предметы быта, такие как:

- стулья;
- столы;
- системные блоки;
- мониторы;

- мультимедийное оборудование больших габаритов;
- вентиляторы;
- радиаторы отопления.

При работе с описанными выше предметами, необходимо соблюдать простые инструкции:

- не пытаться починить оборудование без соответствующих знаний;
- не переставлять мебель и оборудование без согласования с руководством;
- отключать оборудование перед перемещением на другую позицию.

Термические опасности в офисе могут быть связаны со следующими причинами:

- радиаторы отопления;
- источники бесперебойного питания ПК;
- компьютеры.

Для того, чтобы бороться с термическими опасностями нужно соблюдать следующие правила:

- не сливать воду с радиаторов отопления;
- не разбирать любую технику во включенном состоянии;
- осторожно пользоваться приборами быта при разогреве пищи и приготовления напитков с горячей водой.

Электробезопасность является опасным фактором и обычно она связана со следующими источниками:

- поражение электрическим током;
- статическое электричество;
- молниезащита.

Мероприятия защиты при электробезопасности следующие:

- отключать электрооборудование при его ремонте;
- периодически снимать электростатическое напряжение, касаясь пальцами рук, заземленных поверхностей;
- для безопасности во время гроз необходимо удостовериться о наличие молниеотвода, и того факта что все розетки в офисном помещении заземлены.

Пожаровзрывобезопасность характеризуется следующими причинами:

- возгорание на рабочем месте в связи с неправильным обращением с огнем;
- возгорание на рабочем месте в связи с коротким замыканием.

Помещение, в котором проводится работа, оборудовано датчиками дыма и противопожарной сигнализацией.

Помещение оснащено средствами пожаротушения в соответствии с нормами. На 100 м<sup>2</sup> пола имеется:

- пенный огнетушитель ОП-10 – 1 шт.;
- углекислотный огнетушитель ОУ-5 – 1 шт.;
- ящик с песком на 0,5 м<sup>3</sup> – 1 шт.;
- железные лопаты – 2 шт.

При невозможности самостоятельно потушить пожар необходимо вызвать пожарную команду, после чего поставить в известность о случившемся инженера по технике безопасности.

Офис должен постоянно содержаться в чистоте, ежедневно моется пол, выбрасывается мусор и протирается пыль. Офис обеспечен средствами пожаротушения и сигнализацией о наличии продуктов горения в помещении. Компьютерное оборудование для работы в офисе исправно. Пожарные гидранты, пожарный водопровод и средства пожаротушения исправны и находятся на своих штатных местах в состоянии готовности к работе.

Офисное помещение относится к категории В, согласно СП 12.13130.2009 в зимнее время гидранты утеплены, пожарный водопровод заизолирован и утеплен и не разморожен [26].

В офисе приказом назначается лицо, отвечающее за соблюдение правил пожарной безопасности, за исправное состояние пожарного инвентаря и за применение первичных способов пожаротушения.

Краны противопожарного водопровода оснащены брезентовыми шлангами с брандспойтами. Соединительные головки кранов и шлангов должны иметь резиновые прокладки. Скрученные прорезиненные шланги и брандспойты хранятся в опломбированных шкафчиках, размещенных вблизи кранов.

Ящики и щиты, где хранится противопожарный инвентарь, ручки лопат и пожарных топоров, окрашены в красный цвет, а металлические части периодически смазываются и очищаются для предотвращения коррозии.

## **5.2. Региональная безопасность.**

Научно-технический прогресс не стоит на месте. В его условиях, одной из самых актуальных проблем стала охрана окружающей среды. Именно от окружающей среды и её

охраны зависит качество жизни людей, животных и растений. В случае нарушения правил охраны и защиты окружающей среды применяются штрафные санкции.

При рассмотрении влияния персонального компьютера на атмосферу и гидросферу особо вредных выбросов не выявлено. Анализ воздействия на литосферу сводится к обычному бытовому мусору и отбросам жизнедеятельности человека. В случае выхода из строя компьютеров, они списываются и отправляются на специальный склад, который при необходимости принимает определенные меры по утилизации списанной техники и комплектующих.

### **5.3. Организационные мероприятия обеспечения безопасности**

В соответствии с государственными стандартами и правовыми нормами обеспечения безопасности предусмотрена рациональная организация труда в течение смены, которая предусматривает:

- длительность рабочей смены не более 8 часов;
- установление двух регламентируемых перерывов (не менее 20 минут после 1-2 часов работы, не менее 30 минут после 2 часов работы);
- обеденный перерыв не менее 40 минут.

Обязательно предусмотрен предварительный медосмотр при приеме на работу и периодические медосмотры.

Каждый сотрудник обязан пройти инструктаж по технике безопасности перед приемом на работу и в дальнейшем, должен быть пройден инструктаж по электробезопасности и охране труда.

Предприятие обеспечивает рабочий персонал всеми необходимыми средствами индивидуальной защиты.

Оплата труда, социальные пособия, дополнительные выплаты устанавливаются в соответствии со степенью вредности и опасности выполняемых обязанностей.

Согласно требованиям СанПиН 2.2.2/2.4.1340–03, расстояние между рабочими столами с видеомониторами, равно 2 м, а расстояние между боковыми поверхностями видеомониторов примерно 1,2 м [27]. Площадь на одно рабочее место пользователей персонального компьютера с монитором на базе плоских дискретных экранов (жидкокристаллические, плазменные) - 4,5 м<sup>2</sup>.

### **5.4. Особенности законодательного регулирования проектных решений.**

В сфере информационных технологий не предусмотрено применение режима сокращённого рабочего дня, запрещение использования труда женщин и подростков, наличие рабочих мест с вредными и (или) опасными условиями труда и виды компенсаций за таковой, привлечение трудящихся к работам в ночное время и к сменной

работе; применение спецодежды и СИЗ, применение спецпитания и особого лечебно – профилактического обслуживания, особенности обязательного социального страхования и пенсионного обслуживания.

Отсутствие данных требований обусловлено тем, что разработка программного обеспечения и работа с ним не являются узкоспециализированной деятельностью и не относятся к категории тяжелых работ, а так же не предусматривает ненормированного рабочего графика и привлечение к работе лиц, не достигших совершеннолетия.

В соответствии с законодательством об охране окружающей среды, а также на базе Единой государственной системы предупреждения и ликвидации ЧС в данной сфере работ особенностей законодательного регулирования проектных решений нет.

#### **5.5. Безопасность в чрезвычайных ситуациях.**

На таком объекте как офисное помещение могут возникнуть такие чрезвычайные ситуации (ЧС) как:

- техногенные;
- экологические;
- природные.

Рассмотрим такую ЧС, как пожар в офисе. Эта ЧС может произойти в случае повреждения или замыкания электропроводки, оборудования, обрыву проводов, не соблюдению мер пожаробезопасности в офисе и т.д [28].

Для того, чтобы избежать возникновения пожара необходимо проводить следующие профилактические работы, направленные на устранение возможных источников возникновения пожара:

- проведение инструктажа офисных работников о пожаробезопасности;
- периодическая проверка проводки.

Для того что бы увеличить устойчивость офисного помещения к ЧС необходимо устанавливать системы противопожарной сигнализации, реагирующие на дым и другие продукты горения, а также рекомендуется установка огнетушителей и инструктаж рабочих о плане эвакуации из офиса, а так же назначение ответственных за эти мероприятия. Периодически необходимо проводить ложные тревоги, для проверки готовности работников офиса к ЧС. В офисном помещении присутствуют системы, сигнализирующие о наличие пожара или задымленности помещения, находятся огнетушителей и средств тушения пожара (ведра, лопаты и песок, находящиеся в специально оборудованном шкафу, окрашенному в красный цвет). Также, ответственные

за пожарную безопасность и охрану труда, периодически проводятся инструктажи и учебные тревоги.

В случае возникновения ЧС как пожар, необходимо предпринять меры по эвакуации персонала из офисного помещения в соответствии с планом эвакуации (рисунок 5.1). При отсутствии прямых угроз здоровью и жизни произвести попытку тушения возникшего возгорания огнетушителем [29]. В случае потери контроля над пожаром, необходимо незамедлительно эвакуироваться вслед за сотрудниками по плану эвакуации и ждать приезда пожарной службы. При возникновении пожара должна сработать система пожаротушения, издав предупредительные сигналы, и передав на пункт пожарной станции сигнал о ЧС, в случае если система не сработала, необходимо самостоятельно произвести вызов пожарной службы по телефону 101, сообщить место возникновения ЧС и ожидать приезда специалистов [30 - 32].

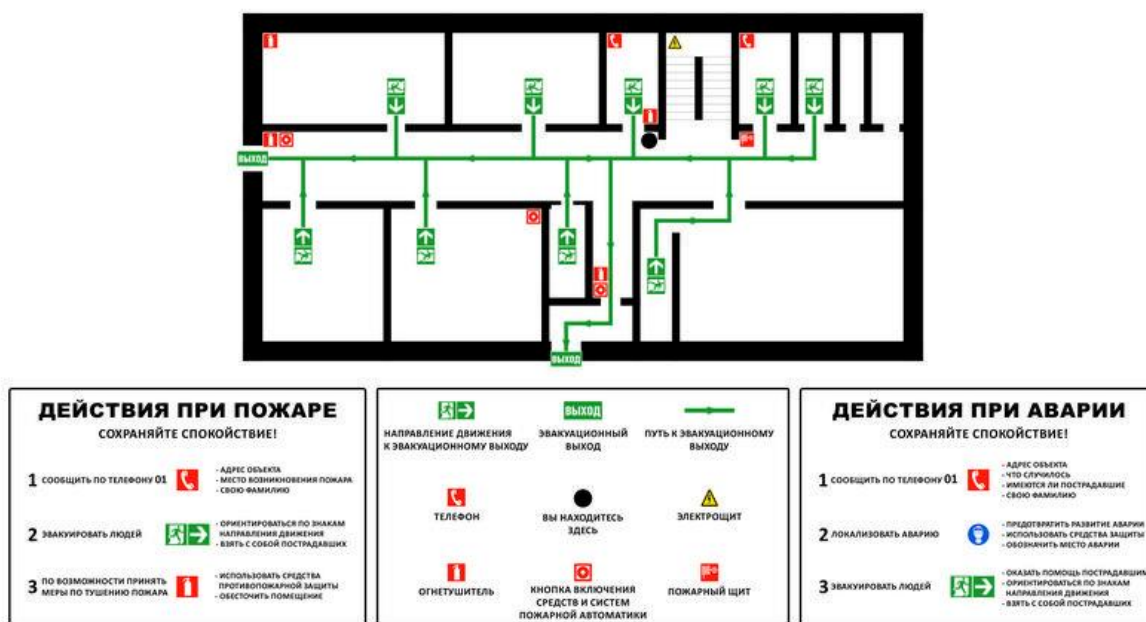


Рисунок 5.1 - План эвакуации

**Выводы по разделу:** разработка данной системы не оказывает вредного воздействия на общество и окружающую среду, в связи соответствием с нормами техногенной безопасности.

## **Заключение**

В результате выполнения выпускной квалификационной работы, была создана система поддержки медицинских научных исследований бронхиальной астмы, на основе анализа клинико-лабораторных показателей.

С использованием разработанной системы было подтверждено предположение о наличии 4 групп (кластеров) пациентов, страдающих БА со сходными значениями клинико-лабораторных показателей.

В рамках выполнения ВКР был разработан алгоритм выявления скрытых закономерностей, который основывается на системе поддержки медицинских исследований БА.

Исследования проводились в рамках выполнения проекта РФФИ 14-07-00675 «Система поддержки медицинских научных исследований психогенных форм бронхиальной астмы» (под рук. профессора каф. ПИ О.Г. Берестневой).

Результаты работы представлены в следующих научных публикациях и выступлениях на конференциях:

1. Жаркова О. С., Шаропин К. А., Сеидова А. С., Берестнева Е. В., Осадчая И. А. Построение систем поддержки принятия решений в медицине на основе деревьев решений // Современные наукоемкие технологии. - 2016 - №. 6-1. - С. 33-37.

2. Сеидова А. С. Инструментальные методы анализа медицинских данных // Научная сессия ТУСУР-2017: материалы Международной научно-технической конференции студентов, аспирантов и молодых ученых, посвященной 55-летию ТУСУРа. В 8 частях, Томск, 10-12 Мая 2017. - Томск: ТУСУР, 2017 - Т. 5 - С. 216-218.

3. Сеидова А. С., Берестнева Е. В. Построение СППР на основе деревьев решений в медицинских исследованиях [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых. В 2 т., Томск, 7-11 Ноября 2016. - Томск: ТПУ, 2017 - Т. 2 - С. 288-289. - Режим доступа: [http://portal.tpu.ru:7777/f\\_ic/files/science/activities/msit/msit2016/Sbornik\\_2016/Sbornik\\_MSIT\\_2016\(Tom2\).pdf](http://portal.tpu.ru:7777/f_ic/files/science/activities/msit/msit2016/Sbornik_2016/Sbornik_MSIT_2016(Tom2).pdf). (Выступление с докладом).

4. Сеидова А. С. Формирование базы знаний для медицинской информационной системы // Научная сессия ТУСУР-2016: материалы Международной научно-технической конференции студентов, аспирантов и молодых ученых. В 6 частях. Часть 3, Томск, 25-27 Мая 2016. - Томск: В-Спектр, 2016 - С. 180-183. (Выступление с докладом).

5. Сеидова А. С., Марухина О. В. Использование пакета WIZWHY для формирования базы знаний медицинской экспертной системы // Информационные

системы и технологии (ИСТ–2016): материалы XXII Международной научно-технической конференции, Нижний Новгород, 22-23 Апреля 2016. - Нижний Новгород: Изд-во НГТУ, 2016 - С. 413.

6. Сеидова А. С. , Берестнева Е. В. , Осадчая И. А. Построение СППР в медицине на основе деревьев решений // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции. В 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 669-671. (Выступление с докладом, диплом 2 степени.).

7. Сеидова А. С. , Марухина О. В. Использование пакета WIZWHY для формирования базы знаний экспертных систем // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции. В 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 495-497. (Выступление с докладом).

8. Сеидова А. С. Формирование базы знаний на основе пакета WIZWHY для медицинской информационной системы // Молодежь и современные информационные технологии: сборник трудов XIII Международной научно - практической конференции студентов, аспирантов и молодых ученых , Томск, 9-13 Ноября 2015. - Томск: ТПУ, 2016 - Т. 2 - С. 262-263. (Выступление с докладом, диплом 3 степени.).

## **Список публикаций студента**

1. Osipova V. V. , Chudinov I. L. , Seidova A. S. Formalized approach in relational database design // Key Engineering Materials . - 2016 - Vol. 685. - p. 930-933.
2. Жаркова О. С. , Шаропин К. А. , Сеидова А. С. , Берестнева Е. В. , Осадчая И. А. Построение систем поддержки принятия решений в медицине на основе деревьев решений // Современные наукоемкие технологии. - 2016 - №. 6-1. - С. 33-37.
3. Сеидова А. С. Инструментальные методы анализа медицинских данных // Научная сессия ТУСУР-2017: материалы Международной научно-технической конференции студентов, аспирантов и молодых ученых, посвященной 55-летию ТУСУРа. В 8 частях, Томск, 10-12 Мая 2017. - Томск: ТУСУР, 2017 - Т. 5 - С. 216-218.
4. Сеидова А. С. , Берестнева Е. В. Построение СППР на основе деревьев решений в медицинских исследованиях [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых. В 2 т., Томск, 7-11 Ноября 2016. - Томск: ТПУ, 2017 - Т. 2 - С. 288-289.
5. Сеидова А. С. Формирование базы знаний для медицинской информационной системы // Научная сессия ТУСУР–2016: материалы Международной научно-технической конференции студентов, аспирантов и молодых ученых. В 6 частях. Часть 3, Томск, 25-27 Мая 2016. - Томск: В-Спектр, 2016 - С. 180-183.
6. Сеидова А. С. , Марухина О. В. Использование пакета WIZWHY для формирования базы знаний медицинской экспертной системы // Информационные системы и технологии (ИСТ–2016): материалы XXII Международной научно-технической конференции, Нижний Новгород, 22-23 Апреля 2016. - Нижний Новгород: Изд-во НГТУ, 2016 - С. 413.
7. Сеидова А. С. , Берестнева Е. В. , Осадчая И. А. Построение СППР в медицине на основе деревьев решений // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции. В 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 669-671.
8. Сеидова А. С. , Марухина О. В. Использование пакета WIZWHY для формирования базы знаний экспертных систем // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции. В 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 495-497.
9. Сеидова А. С. , Сухоплюева В. С. , Ротарь В. Г. Этапы и тенденции корпоративного управления в России // Информационные технологии в науке,

управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции. В 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 402-403.

10. Румянцева Е. А. , Сеидова А. С. Информационная система сбора, подготовки и анализа документов для проведения итоговой государственной аттестации выпускников // Молодежь и современные информационные технологии: сборник трудов XIII Международной научно- практической конференции студентов, аспирантов и молодых ученых , Томск, 9-13 Ноября 2015. - Томск: ТПУ, 2016 - Т. 2 - С. 16-17.

11. Сеидова А. С. Формирование базы знаний на основе пакета WIZWHY для медицинской информационной системы // Молодежь и современные информационные технологии: сборник трудов XIII Международной научно- практической конференции студентов, аспирантов и молодых ученых , Томск, 9-13 Ноября 2015. - Томск: ТПУ, 2016 - Т. 2 - С. 262-263.

12. Сеидова А. С. , Сухоплюева В. С. Не просто кинект // Материалы 53-й Международной научной студенческой конференции МНСК -2015: Информационные технологии, Новосибирск, 11-17 Апреля 2015. - Новосибирск: Новосибирский государственный университет, 2015 - С. 118.

13. Румянцева Е. А. , Сеидова А. С. Проектирование информационной системы для образовательной среды [Электронный ресурс] // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов II Международной конференции, Томск, 19-22 Мая 2015. - Томск: ТПУ, 2015 - С. 750-752.

14. Сеидова А. С., Сухоплюева В. С. Не просто кинект [Электронный ресурс] //сборник трудов 53 - й Международной научной студенческой конференции: Новосибирск, 11-17 апреля 2015. - Новосибирск: НГУ, 2015 - С. 118.

15. Сеидова А. С. , Сухоплюева В. С. Проблема управления неструктурированной информацией [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых: в 2 т., Томск, 12-14 Ноября 2014. - Томск: ТПУ, 2014 - Т. 2 - С. 140-141.

16. Румянцева Е. А. , Сеидова А. С. Интерпретация английского юмора в рецепции разных народов // Коммуникативные аспекты языка и культуры: сборник материалов XIII Международной научно-практической конференции студентов и молодых ученых: в 3 т., Томск, 28-30 Мая 2013. - Томск: ТПУ, 2013 - Т. 1 - С. 16-18.

17.

Сеидова А. С. Использование графических приложений для создания проекта по благоустройству территории дома детского творчества «У Белого озера» г. Томска

[Электронный ресурс] // Вопросы образования и науки: теоретический и методический аспекты: сборник научных трудов по материалам Международной научно-практической конференции в 11 частях, Тамбов, 31 Мая 2014. - Тамбов: Консалтинговая компания Юком, 2014 - Т. 10 - С. 102-104.

18. Сеидова А. С. Слепые фотографии [Электронный ресурс] // Вопросы образования и науки: теоретический и методический аспекты: сборник научных трудов по материалам Международной научно-практической конференции в 11 частях, Тамбов, 31 Мая 2014. - Тамбов: Консалтинговая компания Юком, 2014 - Т. 10 - С. 104-106.

19. Сеидова А. С. , Сухоплюева В. С. Kinect - больше, чем контроллер для Xbox [Электронный ресурс] // Вопросы образования и науки: теоретический и методический аспекты: сборник научных трудов по материалам Международной научно-практической конференции в 11 частях, Тамбов, 31 Мая 2014. - Тамбов: Консалтинговая компания Юком, 2014 - Т. 10 - С. 106-107.

20. Сеидова А. С. Использование графических приложений для создания проекта по благоустройству территории дома детского творчества «у Белого озера» г. Томска [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов X Международной научно-практической конференции студентов, аспирантов и молодых учёных, Томск, 13-16 Ноября 2012. - Томск: Изд-во Томского политехнического университета, 2012 - С. 458-459.

21. Румянцева Е. А. , Сеидова А. С. Проектирование информационной системы для работы вузов при сдаче выпускной квалификационной работы студентами // Технологии Microsoft в теории и практике программирования: сборник трудов XII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых, Томск, 25-26 Марта 2015. - Томск: ТПУ, 2015 - С. 217-218.

22. Румянцева Е. А. , Сеидова А. С. Информационная система сбора и подготовки документов для проведения итоговой государственной аттестации выпускников // Научная сессия ТУСУР–2015: материалы Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых. В пяти частях. Часть 3, Томск, 13-15 Мая 2015. - Томск: В-Спектр, 2015 - С. 296-298.

23. Румянцева Е.А., Сеидова А.С. Информационная система сбора и подготовки документов для проведения итоговой государственной аттестации выпускников// XX Всероссийская научно-техническая конференция студентов, аспирантов и молодых учёных «Научная сессия ТУСУР - 2015»: материалы всероссийской научно-технической конференции студентов, аспирантов и молодых учёных, в 5т., Томск, 13-15 мая 2015 г. - Томск, ТУСУР, 2015 - Т.3 - С. 296-297.

24. Сеидова А. С. Adobe Flash – как способ самореализации // Информационно-телекоммуникационные системы и технологии (ИТСиТ-2014): материалы Всероссийской научно-практической конференции , Кемерово, 16-17 Октября 2014. - Кемерово: КузГТУ, 2014 - С. 292-293.

25. Сеидова А. С. Графические приложения в ландшафтном дизайне // Информационно-телекоммуникационные системы и технологии (ИТСиТ-2014): материалы Всероссийской научно-практической конференции , Кемерово, 16-17 Октября 2014. - Кемерово: КузГТУ, 2014 - С. 290-291.

26. Сеидова А. С. , Сухоплюева В. С. Разработка мобильных игр с помощью Unity3d // Информационно-телекоммуникационные системы и технологии (ИТСиТ-2014): материалы Всероссийской научно-практической конференции , Кемерово, 16-17 Октября 2014. - Кемерово: КузГТУ, 2014 - С. 294-295.

Сеидова А. С. Слепые фотографии // Перспективы развития информационных технологий : труды Всероссийской молодежной научно-практической конференции, Кемерово, 29-30 Мая 2014. - Кемерово: ГОУ КузГТУ, 2014 - С. 129-130.

## Список использованных источников

1. Global strategy for asthma management and prevention // Pharmther URL: [http://pharmther.nuph.edu.ua/wp-content/uploads/2015/10/GINA\\_Report\\_2015\\_Aug11.pdf](http://pharmther.nuph.edu.ua/wp-content/uploads/2015/10/GINA_Report_2015_Aug11.pdf) (дата обращения: 20.02.2017).
2. Авдеев С.Н. Согласованные рекомендации по применению режима единого ингалятора фиксированной комбинации будесонид/формотерол (SMART) в терапии пациентов с бронхиальной астмой /С.Н. Авдеев // Прак. пульмонология. – 2016. – № 1. – С. 2-15.
3. Безруков Н.С., Еремин Е.Л., Ермакова Е.В., Колосов В.П., Перельман Ю.М. Автоматизиро-ванная система «Medical Toolbox» для диагностики бронхиальной астмы по показателям реоэнцефалографии // Информатика и системы управления. – 2006. – №1(11). – С.73-80.
4. Определение и классификация информационных систем // Tsput URL: [http://tsput.ru/res/informat/sist\\_seti\\_fmo/lekci/lekcii-1.html](http://tsput.ru/res/informat/sist_seti_fmo/lekci/lekcii-1.html) (дата обращения: 22.02.2017).
5. Языков К.Г., Немеров Е.В. К вопросу изучения личностных свойств в психофизиологической реактивности больных бронхиальной астмой на аудиовизуальную стимуляцию //Вестник ТГПУ. 2011. № 6(108). С. 134-138.
6. Гаврилова Т.А., Хорошевский В.Ф. – Базы знаний интеллектуальных систем - СПб.: Питер, 2001.-384 с.
7. Дюк В., Самойленко А. Data Mining: учебный курс – СПб: Питер, 2001. – 386 с.
8. Обзор статистических программ // ScienceFiles URL: <http://www.sciencefiles.ru/section/46/> (дата обращения: 26.02.2017).
9. Бурева Н.Н. Многомерный статистический анализ с использованием ППП «Statistica». Учебно-методический материал по программе повышения классификации «Применение программных средств в научных исследованиях и преподавании математики и механики» - Нижний Новгород, 2007. – 112 с.
10. Необходимо ли вам изучать язык R? // Olap URL: <http://www.olap.ru/home.asp?artId=2895> (дата обращения: 15.03.2017).
11. Язык программирования R и его место среди статистических программ // Samoedd URL: <http://samoedd.com/soft/r-introduction> (дата обращения: 15.03.2017).
12. Big data in the cloud // IBM URL: <http://www.ibm.com/developerworks/library/bd-bigdatacloud/index.html> (дата обращения: 15.03.2017).
13. R (язык программирования) // Википедия URL: [https://ru.wikipedia.org/wiki/R\\_\(язык\\_программирования\)](https://ru.wikipedia.org/wiki/R_(язык_программирования)) (дата обращения: 15.02.2017).

14. Кластерный анализ в R: отделяем мух от котлет // Мои волшебные инструменты URL: <http://mymagictools.blogspot.ru/2015/07/r.html> (дата обращения: 25.02.2017).

15. Методы менеджмента качества: Методы поиска новых идей и решений / Е.А. Кузьмина, А.М. Кузьмин. - Т., 2003. - 187 с.

16. Методы менеджмента качества: Функционально-стоимостный анализ. Экскурс в историю/ Е.А. Кузьмина, А.М. Кузьмин. - Т., 2002. - 216 с.

17. Основы функционально-стоимостного анализа: Учебное пособие/ М.Г. Карпунина и Б.И. Майданчика. - М.: Энергия, 1980. - 175 с.

18. Производственная и экологическая безопасность: Методические указания по разработке раздела выпускной квалификационной работы для студентов всех форм обучения/[Электронный ресурс] Сост. М.Э. Гусельников, В.Н. Извеков, Н. В. Крепша, В.Ф. Панин. – Томск.,2010. – 265 с.

19. Микроклимат[Электронный ресурс] / Журнал «Охрана труда и техника безопасности» №5; ред. Корниенко К. И.; Разработчик Гнедов Н. А. - Электрон, дан. - М.: БЖД., 2011. URL: <http://www.otb.by/articles/mikroklimat/> -Загл. с экрана.— Яз.рус. Дата обращения: 26.05.2015 г.

20. Электромагнитные поля на рабочем месте [Электронный ресурс] / Пособие по безопасности жизнедеятельности; ред. Мухина Е. С.; Web-разработчик Плашкевич Е. О. - Электрон, дан. - М.: Электромагнит., 2004. URL: <http://habrahabr.ru/post/140431/>, -Загл. с экрана.— Яз.рус. Дата обращения: 27.05.2015 г.

21. Естественное и искусственное освещение [Электронный ресурс] / Система нормативных документов в строительстве; ред. Журба В. С.; Web-разработчик Зайцева Е. П. - Электрон, дан. - М.: Освещение., 2007. URL: [http://www.tehbez.ru/Docum/DocumShow\\_DocumID\\_312.html/](http://www.tehbez.ru/Docum/DocumShow_DocumID_312.html/) , - Загл. с экрана.— Яз.рус. Дата обращения: 28.05.2015 г.

22. Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки [Электронный ресурс] / Журнал «Охрана труда и техника безопасности» №6; ред. Корниенко К. И.; Разработчик Гнедов Н. А. - Электрон, дан. - М.: БЖД., 2011. URL: <http://www.vashdom.ru/sanpin/224-218562-96/> - Загл. с экрана.— Яз.рус. Дата обращения: 27.05.2015 г.

23. ГОСТ 12.0.003-74. Классификация производственных факторов. – М.: Госстандарт РФ, 1999 г.

24. ГОСТ 12.1.002-84. Система стандартов безопасности труда. Электрические поля промышленной частоты. Допустимые уровни напряженности и требования к проведению контроля на рабочих местах. – М: Стандартинформ, 2003.

25. СанПиН 2.2.4.1191-03. Санитарно-эпидемиологические правила и нормативы «Электромагнитные поля в производственных условиях». – М.: Госкомсанэпиднадзор России, 2003.

26. СанПиН 2.6.1.2523-09. Нормы радиационной безопасности (НРБ-99/2009). – М.: Госкомсанэпиднадзор, 2009.

27. СанПиН 2.2.1/2.1.1.1278-03. Гигиенические требования к естественному, искусственному и совмещенному освещению жилых и общественных зданий. – М.: Госкомсанэпиднадзор, 2003.

28. Правила пожарной безопасности в Российской Федерации. – М.: Министерство Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий, 2003.

**Приложение А**  
**(справочное)**  
Исходные данные

| ФИО              | Группа | Возраст | Рост | Вес | МОД1  | МОД2 | ЖЕЛ1  | ЖЕЛ2  | ФЖЕЛ1 | ФЖЕЛ2 | ОФВ1_1 | ОФВ1_2 | ОФВ1/ЖЕЛ1 | ОФВ1/ЖЕЛ2 | МВЛ1  | МВЛ2  | ПОС1 |
|------------------|--------|---------|------|-----|-------|------|-------|-------|-------|-------|--------|--------|-----------|-----------|-------|-------|------|
| Хусанова Ю.В.    | ВАР1   | 18      | 168  | 55  | 11,5  | 9    | 93    | 95,1  | 76,1  | 88,5  | 71,4   | 82,6   | 76,7      | 84,3      | 49    | 67,9  | 58,6 |
| Псарева Н.Е.     | ВАР1   | 48      | 154  | 70  | 10,3  | 7,7  | 79,1  | 83,61 | 77,2  | 92,3  | 76,6   | 86,6   | 96,19     | 102,18    | 55,3  | 62,2  | 40,1 |
| Аношина М.В      | ВАР1   | 31      | 170  | 58  | 8,2   | 8,5  | 77,1  | 90,5  | 84,4  | 87,3  | 67,5   | 84,4   | 79,37     | 93,01     | 55,5  | 76,7  | 65,8 |
| Газетова Е.А.    | ВАР1   | 36      | 160  | 62  | 7,9   | 7,7  | 81,1  | 85,5  | 67,7  | 84,6  | 62,4   | 86,8   | 76,6      | 91,14     | 58,7  | 78,14 | 61,9 |
| Шимидт Н.А.      | ВАР1   | 40      | 167  | 85  | 11,4  | 12,7 | 76,4  | 97,8  | 75,4  | 88    | 76,8   | 104,7  | 91,22     | 104,6     | 77,9  | 84,6  | 71,1 |
| Соколова Т.А     | ВАР1   | 48      | 158  | 58  | 10    | 9,6  | 52,5  | 50,5  | 51,3  | 55    | 53,13  | 58,13  | 81,3      | 83,3      | 38,13 | 43,3  | 31,3 |
| Катаев О.В       | ВАР1   | 32      | 181  | 66  | 8,8   | 8    | 57,2  | 65,13 | 50,8  | 72,8  | 50     | 75,2   | 85,6      | 104,88    | 71    | 90,4  | 32,1 |
| Изотова Е.И      | ВАР1   | 41      | 164  | 80  | 9     | 6,5  | 60,17 | 64,5  | 54,64 | 72,5  | 51,7   | 62,08  | 76,2      | 84,3      | 43,06 | 51,3  | 36,9 |
| Яткина Г.Н       | ВАР1   | 49      | 156  | 82  | 7,5   | 7,7  | 82,1  | 82    | 96,6  | 106,6 | 104,8  | 112,9  | 101,2     | 112       | 108,4 | 113,5 | 84,8 |
| Шумилина Л.В.    | ВАР1   | 43      | 156  | 52  | 9,32  | 8,92 | 101   | 102   | 98,7  | 101   | 83,6   | 96,7   | 82,2      | 88,1      | 85,9  | 90,2  | 60,8 |
| Василинин А.Н.   | ВАР1   | 45      | 176  | 84  | 5,8   | 4,6  | 88,6  | 94,2  | 81    | 90,8  | 78,2   | 82,4   | 81        | 87,1      | 68    | 75,7  | 66,4 |
| Давыдова Н.М.    | ВАР1   | 42      | 167  | 84  | 9,03  | 6,97 | 72,42 | 78    | 74,2  | 76    | 57,5   | 68,5   | 79,37     | 86,71     | 39,9  | 46,84 | 38,9 |
| Тайчинов А.В.    | ВАР1   | 38      | 160  | 56  | 8,99  | 7,03 | 104,7 | 107   | 83,9  | 76,5  | 75     | 64,9   | 71        | 82,5      | 104   | 79,6  | 63,8 |
| Старкова В.В.    | ВАР1   | 47      | 150  | 80  | 7,5   | 7,7  | 73,4  | 80,9  | 73    | 81,2  | 62,7   | 85,5   | 81        | 87,1      | 73,2  | 101,6 | 67,5 |
| Голina С.А.      | ВАР1   | 41      | 164  | 71  | 9,4   | 8,4  | 63,04 | 75,9  | 68,75 | 85,12 | 57,3   | 76,39  | 82,44     | 91,26     | 51,96 | 52    | 44,1 |
| Хаджаева А.А.    | ВАР1   | 16      | 160  | 50  | 10,66 | 4,5  | 57,8  | 76,12 | 53,5  | 75,1  | 58,8   | 78,1   | 84,5      | 102,56    | 36,3  | 40,7  | 45,8 |
| Козлов С.П.      | ВАР1   | 42      | 162  | 62  | 6     | 6,59 | 107,1 | 110,2 | 79,3  | 98,7  | 70,8   | 98     | 66,8      | 88,4      | 81,9  | 73,4  | 53,2 |
| Зарубина Л.Ф.    | ВАР1   | 46      | 163  | 64  | 8,5   | 8,6  | 71,3  | 77,1  | 105,3 | 122,7 | 106,9  | 124,5  | 88,8      | 101,9     | 102,2 | 118,9 | 76,6 |
| Гибадуллина Ф.В. | ВАР1   | 52      | 168  | 77  | 6,22  | 6,25 | 55,2  | 64,8  | 56,1  | 58,1  | 55,7   | 59,1   | 56,1      | 63,6      | 42,1  | 48,6  | 24,1 |
| Севастьянов В.А. | ВАР1   | 49      | 178  | 100 | 8,52  | 7,5  | 88,4  | 90,1  | 77,4  | 96,2  | 75,7   | 92,2   | 72,7      | 84,6      | 75,8  | 86,4  | 43,7 |
| Ревичев В.М.     | ВАР1   | 43      | 172  | 70  | 8,2   | 7,1  | 80    | 83    | 59,2  | 71,1  | 59,3   | 71     | 72        | 84        | 58,1  | 56,2  | 45   |
| Донцова С.В.     | ВАР1   | 48      | 158  | 60  | 8,3   | 7,7  | 83    | 85,5  | 83    | 91    | 76     | 89     | 82,1      | 86,3      | 78    | 86    | 62,3 |
| Саранцева С.И.   | ВАР1   | 38      | 144  | 34  | 8,6   | 8,1  | 75,5  | 66,1  | 75,1  | 80    | 65     | 72     | 72        | 76        | 73    | 75    | 46,2 |
| Афанасьев П.И.   | ВАР1   | 40      | 168  | 85  | 8,1   | 7,4  | 100   | 101   | 100,1 | 106,6 | 79,1   | 86,6   | 79        | 86        | 79,3  | 88,1  | 66,1 |



| ФИО              | Группа | ПОС2        | МОС25_1     | МОС25_2     | МОС50_1     | МОС50_2     | МОС75_1     | МОС75_2     | Сdyn1 | Сdyn2 | Сstat1 | Сstat2 | Ran1 | Ran2 | Ranr1 | Ranr2 | Wobm1       | Wobm2 |
|------------------|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------|-------|--------|--------|------|------|-------|-------|-------------|-------|
| Хусанова Ю.В.    | ВАР    | 83,3        | 61,3        | 88,4        | 59,8        | 70,3        | 47,9        | 63,9        | 0,125 | 0,13  | 0,18   | 0,19   | 3,8  | 2,3  | 7,2   | 4,1   | 0,469       | 0,27  |
| Псарев Н.Е.      | ВАР    | 43,2        | 45,1        | 47,1        | 46,9        | 48,1        | 46,3        | 51,2        | 0,06  | 0,05  | 0,1    | 0,09   | 4,6  | 4,8  | 6,5   | 6,4   | 0,594       | 0,27  |
| Аношина М.В.     | ВАР    | 94,5        | 60,2        | 86,3        | 58,9        | 81,1        | 48,2        | 63,6        | 0,17  | 0,225 | 0,19   | 0,22   | 5,5  | 2,1  | 5,6   | 2,1   | 0,2         | 0,159 |
| Газегова Е.А.    | ВАР    | 82,5        | 58,2        | 89,1        | 41,8        | 65,6        | 41,7        | 58,2        | 0,127 | 0,2   | 0,138  | 0,22   | 5,52 | 3,1  | 6,5   | 3,2   | 0,285       | 0,137 |
| Шмидт Н.А.       | ВАР    | 89,2        | 51,6        | 81,2        | 45,1        | 66,5        | 40,2        | 57,1        | 0,036 | 0,091 | 0,17   | 0,16   | 5,2  | 2,85 | 8,12  | 4,65  | 0,488       | 0,306 |
| Соколова Т.А.    | ВАР    | 30,2        | 38,11       | 37,2        | 22,2        | 18,1        | 23,9        | 30,8        | 0,08  | 0,07  | 0,1    | 0,11   | 6,46 | 3,55 | 8,4   | 4,75  | 0,938       | 0,45  |
| Кагаев О.В.      | ВАР    | 47,4        | 38,1        | 47          | 37,4        | 46,6        | 29,2        | 31,1        | 0,15  | 0,18  | 0,16   | 0,19   | 4,8  | 3,1  | 5,3   | 4,1   | 0,792       | 0,54  |
| Изотова Е.И.     | ВАР    | 53,8        | 44,1        | 60,34       | 42,4        | 60,6        | 34          | 48,08       | 0,05  | 0,065 | 0,1    | 0,11   | 5,38 | 3,75 | 8,45  | 5,52  | 1,125       | 0,65  |
| Яткина Г.Н.      | ВАР    | <b>98,3</b> | <b>83,8</b> | <b>89,6</b> | <b>69,6</b> | <b>72,3</b> | <b>65,1</b> | <b>77,7</b> | 0,07  | 0,144 | 0,12   | 0,14   | 5,29 | 4,5  | 6,5   | 4,5   | 0,469       | 0,371 |
| Шумилина Л.В.    | ВАР    | 70,2        | 56,6        | 65,1        | 54,9        | 62,1        | 46,6        | 31,2        | 0,09  | 0,12  | 0,1    | 0,14   | 5,1  | 3,29 | 8,5   | 4,9   | 0,678       | 0,41  |
| Василишин А.Н.   | ВАР    | 81,2        | 55,1        | 61,1        | 51          | 60,3        | 41,8        | 50,2        | 0,06  | 0,08  | 0,1    | 0,082  | 4,41 | 3,25 | 6,4   | 4,27  | 0,365       | 0,238 |
| Давыдова Н.М.    | ВАР    | 58,3        | 37,2        | 59,6        | 23,2        | 35,1        | 33,5        | 50,3        | 0,1   | 0,09  | 0,15   | 0,12   | 6,3  | 3,3  | 9,1   | 4,71  | 0,749       | 0,632 |
| Тайчинов А.В.    | ВАР    | 58          | 61,4        | 55,5        | 51,6        | 41,4        | 41,8        | 30,7        | 0,105 | 0,11  | 0,112  | 0,12   | 4,32 | 3,61 | 8,62  | 4,89  | 0,504       | 0,34  |
| Старкова В.В.    | ВАР    | 83,6        | 68          | 88          | 50,2        | 64,4        | 49,4        | 61,2        | 0,077 | 0,129 | 0,11   | 0,14   | 4,2  | 2,73 | 5,9   | 3,27  | 0,375       | 0,159 |
| Голина С.А.      | ВАР    | 60,87       | 38,45       | 44,94       | 32,96       | 45,89       | 15,87       | 19,04       | 0,077 | 0,08  | 0,11   | 0,09   | 6,5  | 3,7  | 8,6   | 5,75  | 0,65        | 0,526 |
| Хаджаева А.А.    | ВАР    | 59,7        | 49,3        | 61,8        | 34,7        | 38,6        | 24,9        | 18,8        | 0,056 | 0,068 | 0,128  | 0,085  | 5,11 | 2,67 | 6,69  | 2,82  | 0,661       | 0,126 |
| Козлов С.П.      | ВАР    | 70,1        | 43,2        | 62,5        | 39,6        | 56,5        | 39,5        | 46,2        | 0,157 | 0,145 | 0,186  | 0,174  | 5,61 | 3,38 | 7,73  | 3,8   | 0,371       | 0,225 |
| Зарубина Л.Ф.    | ВАР    | 100,1       | 58,1        | 80,1        | 76,4        | 96,5        | 56,1        | 76,6        | 0,07  | 0,08  | 0,06   | 0,1    | 5,1  | 4,1  | 7,1   | 5,14  | 0,966       | 0,486 |
| Гибатуллина Ф.В. | ВАР    | 23,9        | 24,7        | 23,9        | 22,2        | 18,5        | 15,3        | 16,7        | 0,08  | 0,07  | 0,1    | 0,09   | 5,8  | 4,9  | 8,8   | 5,2   | 0,8         | 0,68  |
| Севастьянов В.А. | ВАР    | 56,9        | 43,9        | 52,9        | 40,3        | 52          | 54,6        | 68,7        | 0,136 | 0,11  | 0,146  | 0,126  | 5,45 | 3,3  | 7,55  | 3,33  | 0,504       | 0,29  |
| Ревнев В.М.      | ВАР    | 61          | 44,1        | 56,2        | 46          | 52          | 38,1        | 49,3        | 0,18  | 0,17  | 0,19   | 0,18   | 5,1  | 3,8  | 6,2   | 3,95  | 0,61        | 0,48  |
| Донцова С.В.     | ВАР    | 63,3        | 66,1        | 67,5        | 52,4        | 66,2        | 43          | 51,4        | 0,1   | 0,11  | 0,115  | 0,138  | 4,21 | 3,98 | 6,63  | 4,31  | 0,61        | 0,474 |
| Сарапцева С.И.   | ВАР    | 58,9        | 45          | 68          | 38,1        | 50,2        | 45          | 54          | 0,086 | 0,087 | 0,11   | 0,112  | 3,6  | 1,8  | 5,2   | 3,1   | 0,78        | 0,41  |
| Афанасьев П.И.   | ВАР    | <b>68,1</b> | 61,4        | <b>65,2</b> | 51,6        | 60,8        | 41,8        | <b>52,1</b> | 0,106 | 0,105 | 0,112  | 0,11   | 4,9  | 3,6  | 6,82  | 4,87  | <b>0,55</b> | 0,34  |

**Приложение Б**  
**(справочное)**

**Разделы:**

Введение

Глава 1. Постановка задачи

Глава 2. Выбор решения задачи

Студент:

| Группа | ФИО                        | Подпись | Дата |
|--------|----------------------------|---------|------|
| 8KM51  | Сеидова Айсель Султан кызы |         |      |

Консультант кафедры ПИ:

| Должность      | ФИО             | Ученая степень,<br>звание | Подпись | Дата |
|----------------|-----------------|---------------------------|---------|------|
| Доцент каф. ПИ | Чердынцев Е. С. | К.Т.Н.                    |         |      |

Консультант – лингвист кафедры ИЯ:

| Должность                      | ФИО             | Ученая степень,<br>звание | Подпись | Дата |
|--------------------------------|-----------------|---------------------------|---------|------|
| Зав.каф. Иностранных<br>языков | Сидоренко Т. В. | К.П.Н                     |         |      |

## **Introduction**

The Bronchial Asthma (BA) is a global problem, which relevance increases throughout the world. Bronchial asthma is called one of the illnesses of civilization as its prevalence in the modern society will constantly increase. The statistics demonstrates that the disease at children's age in 60–80% of cases proceeds at adult patients.

The Fast growth of an asthma case rate asthma happened at the end of the 20th century. In 1998 the amount of patients with asthma in the world equaled about 155 million people, and now this number has reached 300 million. According to some prognostic analytical research, by 2025 in case if the process of urbanization will continue at the same rate, bronchial asthma will infect 100–150 million people [1] .

According to the statistical data of 2014, in Russia 1406493 patients with bronchial asthma are officially registered. By experts forecasting, the amount of patients with asthma will exceed the official statistics at least by 5–6 times and make/ reach about 9,915 million [2].

The medical data, which are saved up in the archives, are urgent for studying this disease, largely due to a huge stock of information on various cases of diseases, the course and gravity of treatment, value of various clinical and laboratory indicators.

According to the Federal Information Fund, the data of social and hygienic monitoring (2014) the Tomsk region is called as the risk zone. In 2014 the asthma case rate (the asthmatic status) of the population in the region was in group age at the following levels: children – 238,2 on 100 thousand population of this group age group, teenagers – 142,9, adults – 107,8.

In this regard, the relevance to study this within the final qualification work is obvious. The analysis of the clinical laboratory indicators reflecting a state and a clinical picture and finding of the patterns helping to correct treatment can be used to find a solution to this problem.

The purpose of the work is to create the information system of data processing by using the data mining methods.

The final qualification work is devoted to ways helping to find the solution for the following tasks:

- to reveal the hidden regularities of efficiency of patients treatment of;
- to create groups of patients with similar values of clinical laboratory indicators;
- to create the knowledge base on the basis of hidden regularities;
- to create a structure of the information system and to choose tools for it to be implemented.

## **Chapter 1.**

### **Problem definition**

#### **1.1. Concept of an information system**

The Information System (IS) is the interconnected set of the means, methods and personnel used for storage, processing and issue of information for the benefit of achievement of a goal.

The term an information system is used both in loose meaning of the word, and in a narrow sense.

In a loose meaning of the word the information system is the set of technical, program and organizational support, and also a staff intended timely to provide appropriate people with appropriate information [3].

In a narrow sense the information system is called only a subset of the ICs components in a broaden sense, including databases, DBMS and specialized application programs.

## **1.2 Description of subject domain**

City clinical hospital №3 of Tomsk has provided with the data array with the clinical laboratory indicators of patients before and after the treatment. These patients are the adults with bronchial asthma. The solution to the task is knowledge database creation using the available indicators. The knowledge base will be used further in an information support system of decision-making when performing treatment.

This study will involve 83 patients. The current problem is solved on the ground of the basic data containing the following indicators:

BANP – Bronchial Asthma Not Psychogenic

BASP – Bronchial Asthma Somato — Psychogenic

BAPI – Bronchial Asthma which is Psychogenically Induced

PD – Psychogenic Dyspnea

Group code – designation (digits or letters) of one or another patient group:

Sex – sex of the patient

Age – age of the patient

Human height – body height of the patient

Weight –weight of the patient

DVN — degrees of ventilatory difficulty

SBA — severity of bronchial asthma

Age 1<sup>st</sup>.ep. - — age of the patient during the first illness episode (age)

DI — duration of illness

Kl.pr.atopy – clinical signs of atopy

Indicators of lung ventilation: 1-before, 2 after AVS

PMV – Pulmonary Minute Volume

PC – Pulmonary Capacity, % (to due value)

FPC – Forced Pulmonary Capacity, %

FEV1 – Forced Expiratory Volume, %

FEV1/PC –Ratio of FEV1 to PC, %

MPV –Maximum Pulmonary Ventilation per 1min, %

PEVF. – Peak Expiratory Volume Flow, %

PEVF25 – Peak Expiratory Volume Flow is at the level of 25% of FEV1

PEVF50 – Peak Expiratory Volume Flow is at the level of 50% of FEV1

PEVF75 – Peak Expiratory Volume Flow is at the level of 75% of FEV1

It must be noted that all the indicators of lung ventilation (except PMV) are expressed in % to due values

Pulmonary mechanics indicators: 1-before, 2 after AVS

DLC – Dynamic Lung Compliance (C), l/cmH<sub>2</sub>O

SLC – Static Lung Compliance (C), l/cmH<sub>2</sub>O

ARI — Airways Resistance to Inspiration, l/cmH<sub>2</sub>O/sec

ARE – Airways Resistance to Expiration, l/cmH<sub>2</sub>O/sec

TSBW – Total Spontaneous Breathing Work, kg/m/min

SBW – Specific Breathing Work (per ventilation liter), kg/m/l

NCFTBW – Non-crushing Fraction of Total Breathing Work, kg/m/min

CFTBW – Crushing Fraction of Total Breathing Work, kg/m/min

W PMV10 – Total Breathing Work at PMV10, kg/m/min

W PMV15 – Total Breathing Work at PMV15, kg/m/min

W MPVtotal – Total Breathing Work at MPV, kg/m/min

W MPVspecific – Breathing Work at MPV, kg/m/l

Pulmonary Mechanics Indicators are given further according to the SI system

During performance of work the indicators are considered both before and after the treatment given, and the indicators difference is considered as well.

#### **1.4 Formal assignment of task**

An initial data array of our task is the matrix  $X_{nm}$ ,

where n – the number of patients (n=83);

m – the number of medical parameters (m=48).

The multidimensional analysis is used to complete this task. A subject of the multidimensional analysis is difficult systems that elements are characterized by a set of interdependent objects and signs interdependent, presented in the matrix form which lines correspond to observed objects, and columns – to the sign characterizing them [4].

To complete this task it is necessary to analyze the existing approaches to the analysis of multidimensional data, to choose the method and the way of its computer realization which allows to process big data arrays and to reveal the latent meaning.

### **Chapter 2 Choice of the task completion**

#### **2.1 Choice of technology and the description of methods of the task completion**

##### **2.1.1 Knowledge base**

The knowledge base is the special type of database developed for control of knowledge, as collection, storage, search and output of knowledge. For data storage the databases (the large volume and rather small cost of information are defined for them) are used and for storage of knowledge the knowledge base is used (small volume, but exclusively expensive information arrays).

Knowledge is regularities of data domain (principles, links and laws) received as the result of practical activities and professional experience, allow to someone to do something to put and to solve problems in this area.

##### **2.1.2 Strategy of acquisition of knowledge**

The key issue is the process of acquisition of knowledge when forming the knowledge base. To name of this process in the literature several terms were emulated: acquisition, extraction and formation of knowledge.

Acquisition of knowledge is understood as a way of the automated filling of the knowledge base by the dialogue between the expert and the special program.

Extraction of knowledge is a procedure of interaction of the knowledge engineer with a source of knowledge (the expert, etc.). The procedure of knowledge extraction is time- and labor-consuming as the knowledge engineer (which is equipped with special knowledge of psychology, system analysis, mathematical logic and so forth) should recreate the model of domain area which is used by the experts for decision-making.

The term «knowledge formation» was assigned to extremely advanced and actively developing field of knowledge engineering which is engaged in development of models, methods and algorithms of training [5]. It includes the inductive models of knowledge formation and automatic hypotheses generation on the basis of tutorial selections, training by analogy and other methods. These models allow to reveal cause-and-effect empirical dependences in the databases with incomplete information containing the structured numerical and character objects (it is frequent in the conditions of incompleteness of information).

Knowledge formation is a process of data analysis and detection of the insights with use of a special mathematical apparatus and software.

Therefore, it is possible to select three main methods of acquisition of knowledge:

1. To use a computer with suitable program tools, otherwise acquisition of knowledge;
2. The direct interaction between the knowledge engineer and the source of knowledge (it could be the expert, special literature or other sources) without ADP equipment (whether it be the expert, special literature or other sources), otherwise knowledge extraction;
3. To use the tutorial programs in case of the presence of representative (that is rather representative) selections of examples of decision-making in the domain area and the appropriate application program packages, otherwise knowledge formation.

### **2.1.3 Data mining Technology**

The term data mining is often translated as the data search, excavation or «extraction of knowledge grains from the data mountains». The Data mining technology is defined rather precisely by one of founders of this direction Grigory Piatetsky-Shapiro: «Data Mining is a process of detection in raw data of knowledge earlier unknown, uncommon, practically useful and available to interpretation that are necessary for decision-making in various spheres of human activity».

Data Mining – a set of various methods knowledge detection. The choice of epy method often depends on the type of available data and on what information needs to be obtained.

The data Mining methods and algorithms are following: artificial neural networks, decision trees, symbolical rules, nearest neighbor and the k-nearest neighbor methods, the method of support vectors, Bayesian networks, linear regression, correlation and regression analysis; hierarchical methods of the

cluster analysis, not hierarchical methods of the cluster analysis, including algorithms of k-means and k-median; methods of search of associative rules, including algorithm Apriori; method of limited search, evolutionary programming and genetic algorithms, various methods of data visualization and set of other methods.

After studying abovementioned methods and algorithms, based on experience of our colleagues from laboratory of information technologies in the social sphere and medicine in cooperation with the Tomsk Research Institute of Balneology and Physical Therapy, for the task completion the cluster analysis and the creation of regression trees are used [6].

#### **2.1.4. Cluster analysis**

The clustering is a group of objects on the basis of the data describing the entity of these objects. Objects in the cluster should be «similar» against each other and different from the objects from other clusters. The more objects are similar in the cluster and the more distinctions exist between clusters, the more clustering is precise.

There are different methods of cluster analysis that could be divided into two groups:

##### **1. Hierarchical:**

Hierarchical agglomerative methods (Agglomerative Nesting, AGNES);

Hierarchical divisible methods (DIvisive ANALysis, DIANA).

##### **2. Not hierarchical (iterative) methods:**

- Algorithm of k-means;

- Algorithm of PAM (partitioning around Medoids).

It is possible to receive different decisions for the same data, using different methods of cluster analysis.

The most common algorithm of k-means is applied for this study. The algorithm divides objects into k groups where each group represents one cluster. The fundamental difference of the k-means method from the hierarchical cluster analysis is that it is necessary to define initially number of clusters into which it is required to break the researched set. Respectively, it is desirable before the beginning of the analysis to have the hypothesis of the researched set structure. The structure of medical data was probed and clustering cases for  $k = 4$  were carried out using this method.

#### **2.1.5. Creation decision trees**

Decision trees divide objects spaces according to the some rule set of partition. These rules are logical statements concerning this or that variable and can be true or false. Here there are three key circumstances: a) rules allow to realize sequential dichotomizing data segmentation, b) two objects are considered similar if they appear in the same segment of division, c) on each step of division the amount of information concerning the researched variable (response) increases [7].

Trees of classification and regression are one of the most popular methods of the decision of many practical tasks that are caused by the following reasons:

1. Decision trees allow to receive really easily interpreted models representing rule set type «if ..., then ...». Interpretation is facilitated, including, with the opportunity to present these rules in the form of intuitive tree structure.

2. Decision trees (due to its structure) allow to work with any type variables without any preliminary preparation of these variables for input in model (for example, logarithmation, conversion of variables from categorical to indicated, etc.).

3. There is no need for the researcher to set the form of correlation between response and predictor as it, for example, happens in a case of normal regression models. It appears to be very useful when working with large volume of data with relatively unknown properties.

4. Decision trees, in fact, automatically execute selection of the informative predictors and consider possible interactions. Particularly it makes decision trees the useful instrument of prospecting data analysis.

5. Decision trees can be applied effectively to data with the missing values that is very useful in case of the decision of practical tasks where existence of the passed values is more the rule than exception.

6. Decision trees are well applicable both to the quantitative, and to qualitative dependent variables.

## **2.2. Choice of the decision tool**

### **2.2.1. Scripting language R**

R – is the universal programming language developed for application in such areas as prospecting data analysis, classical statistical tests and high-level graphics. The language was created as similar to the language S developed in Bell Labs and was his alternative realization though there are essential differences between the languages, but generally the code works at language S in the environment R. Initially R has been developed by Ross Ihaka (English Ross Ihaka) and Robert Gentleman (English Robert Gentleman) (the first letter of their names — R); which are the staff of statistical faculty of University of Auckland. R Foundation organization supports and develops the language and the environment.

Thanks to the extensive and continuously extending library of packages, language R takes the leading positions in statistics, data analysis and data production.

### **2.2.2. Language R data types**

All data objects (and therefore variables) in R can be divided into the following classes (i.e. object types):

- ° numeric – objects which contain integer (integer) and the real numbers (double);
- ° logical – logical objects which accept only two values: FALSE (in abbreviated form F) and TRUE (T);
- ° character – character objects (variable values are set in double, or single quotes).

Unlike a vector or a matrix which may contain only one type data, the list of the data frame could contain any data types. It allows effectively, i.e. in one object, to store heterogeneous information. Each component of the list can be a variable, a vector, a matrix, a factor or other list. Besides, these elements

can belong to different types: numbers, lines of characters, Boolean variables. Lists are the most general means of storage of internal information: in particular, results of the majority of statistic analyses in the program R are kept in objects - lists.

The data frame represents the R-object reminding a leaf of the electronic spreadsheet Microsoft Excel in its structure. Each column of the table is the vector containing data of the certain type. At the same time there is the rule according to which all columns should have identical length works (actually, from «point of view» R-data sheet is a special case of the list in which all components - vectors have the identical size). Data sheets are the main class of the R-objects used for data storage. Usually such tables are prepared by means of external applications (the Microsoft Excel program is especially popular and convenient for it) and then are downloaded into the R-environment.

### **2.2.3. Data import in R**

Data import in R-system often causes problems to those who only start working with this program. Nevertheless, there is nothing difficult with it. The most widespread methods of import of data sheets to a work environment of R explicitly will be considered below; however at first it is worth to study preparation rules of the loaded files:

Imported dataful table should not contain the blank cells. If some values for one or another reason are absent, it is necessary to enter NA instead of them.

It is recommended to transform the imported table into the simple text file with one of available extensions. In practice files with extension of .txt in which values of variables 38 tabulation (tab-delimited files), and also files with extension of .csv (comma separated values) in which variable values are separated by commas or other separating character are usually used.

As the first line in the imported table it is recommended to enter titles of columns variables. Such line –is a convenient, but not obligatory entry of the loaded file. If it is absent, then it is necessary to report about it in the command description which will control the file loading (for example, `read.table()` – see below). All following lines in the file (as the first element) may contain row headings (if those are available), after which there are values of each variables available in the variable table.

It is better to appropriate column names of the using the identification rules of R-variables, i.e. to exclude gaps and other special characters (the exception is point and underlining). In order to avoid the problems with to the coding, all text values in the imported files should be created with use of letters of the Latin alphabet.