



МИРЭА - Российский Технологический Университет
Физико-технологический институт

Использование методов распознавания в задачах автоматизированного определения лекарственных препаратов на основе данных пациента

АВТОР:

КОЛОСОВ И.А.

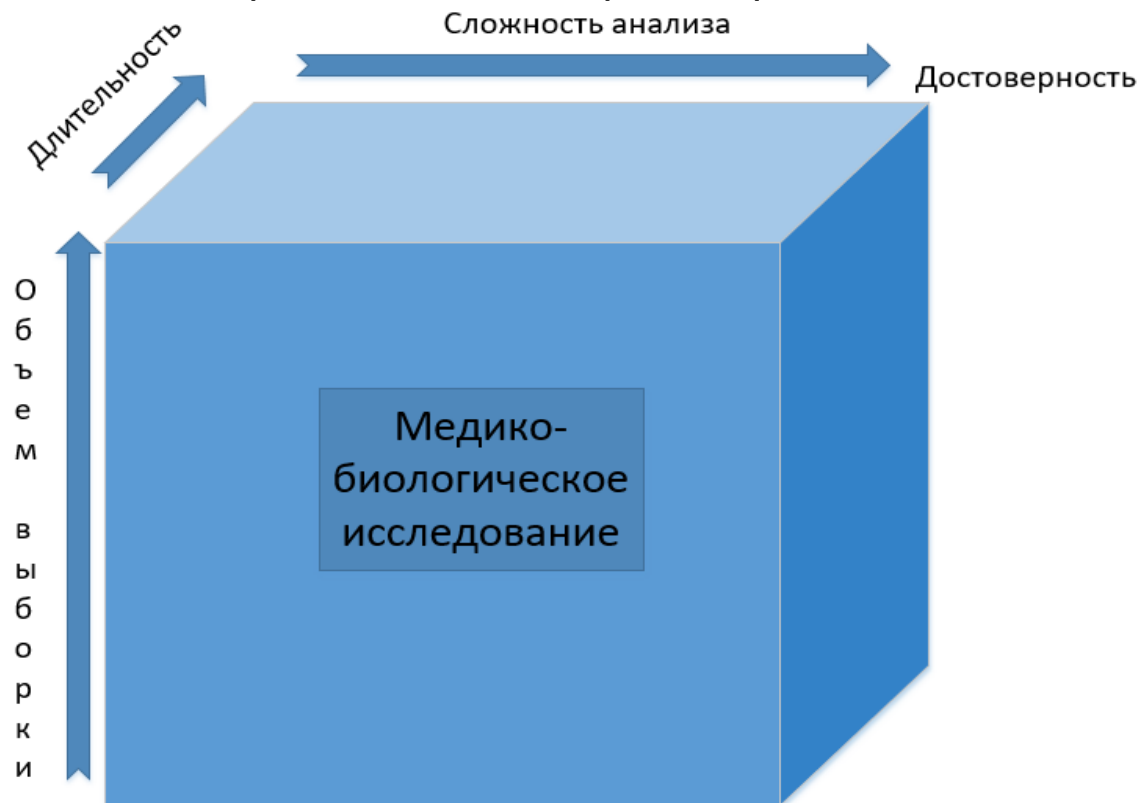
НАУЧНЫЙ РУКОВОДИТЕЛЬ:

К.Т.Н., ДОЦ. МИРЭА СТЕПАНОВ Д.Ю.

Москва 2020

Введение

Проблема: большие временные, финансовые затраты проведения МБИ и внедрения их результатов.



Решение: ведение в лечебную практику сбора данных о действии препаратов на пациентов с различными параметрами и автоматизированный анализ этих данных с использованием методов распознавания с последующим автоматизированным использованием результатов анализа.

Цель и задачи

Цель работы: Использование методов распознавания в задачах автоматизированного определения лекарственных препаратов на основе данных пациента.

Задачи:

1. Изучить итерационную методологию внедрения информационных систем, а также методы распознавания образов.
2. Изучить графические нотации проектирования процессов IDEF0, IDEF3 и данных UML CD.
3. Изучить структуру описания пользовательских программ и виды тестирования программных продуктов.
4. Идентифицировать требования к процессам обучения и тестирования классификатора по предложению лекарственных средств.
5. Спроектировать процессы в IDEF0 и IDEF3 для модели AS-IS и TO-BE до 3-4 уровней детализации; данные – UML CD, включая нормализацию таблиц; структуру приложения, а также блок-схему заданной разрабатываемой функции для каждой итерации.
6. Реализовать и количественно оценить программное приложение для автоматизации работы больницы, включая разработку заданной функции, в среде Python.

1.1 Итерационная модель внедрения

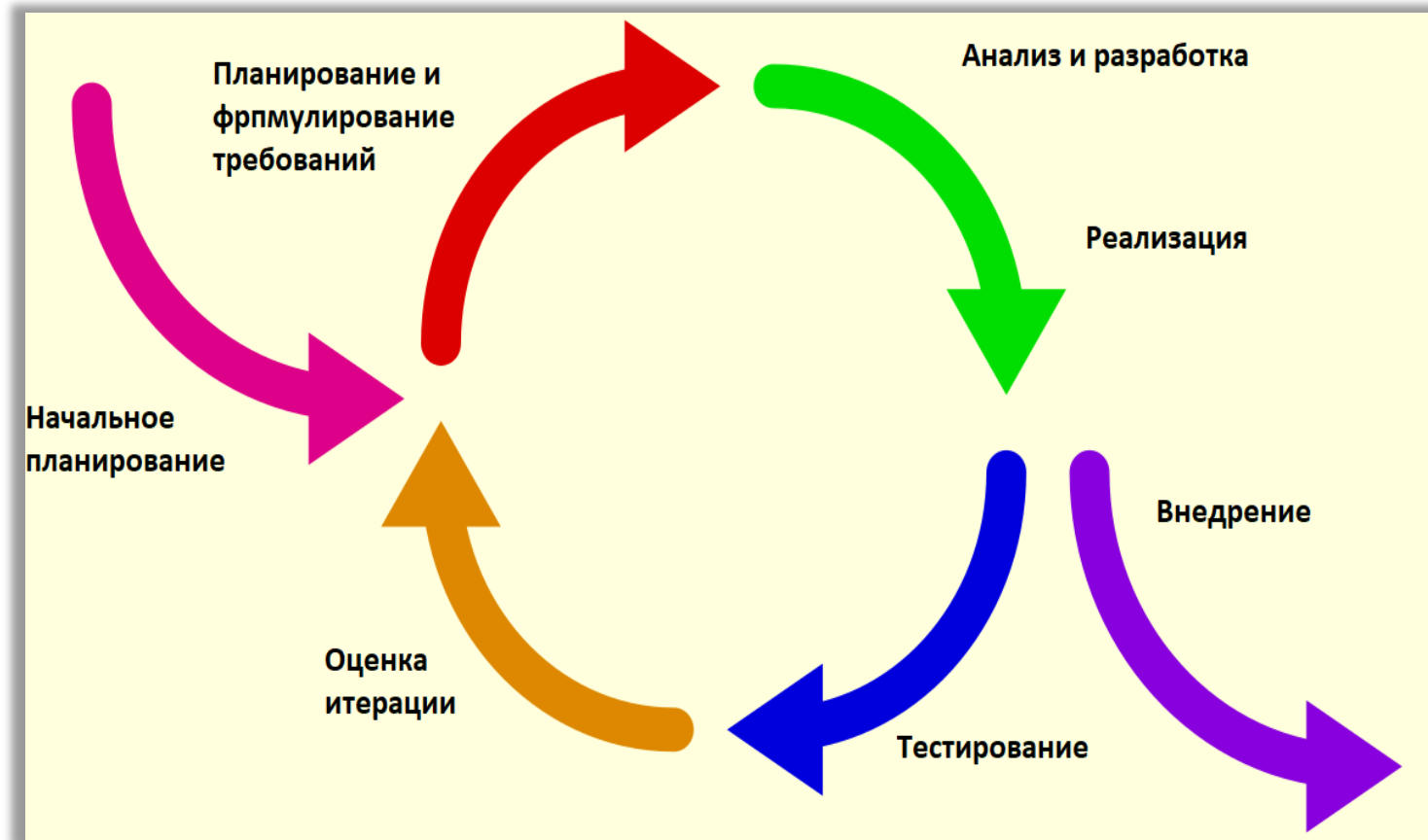
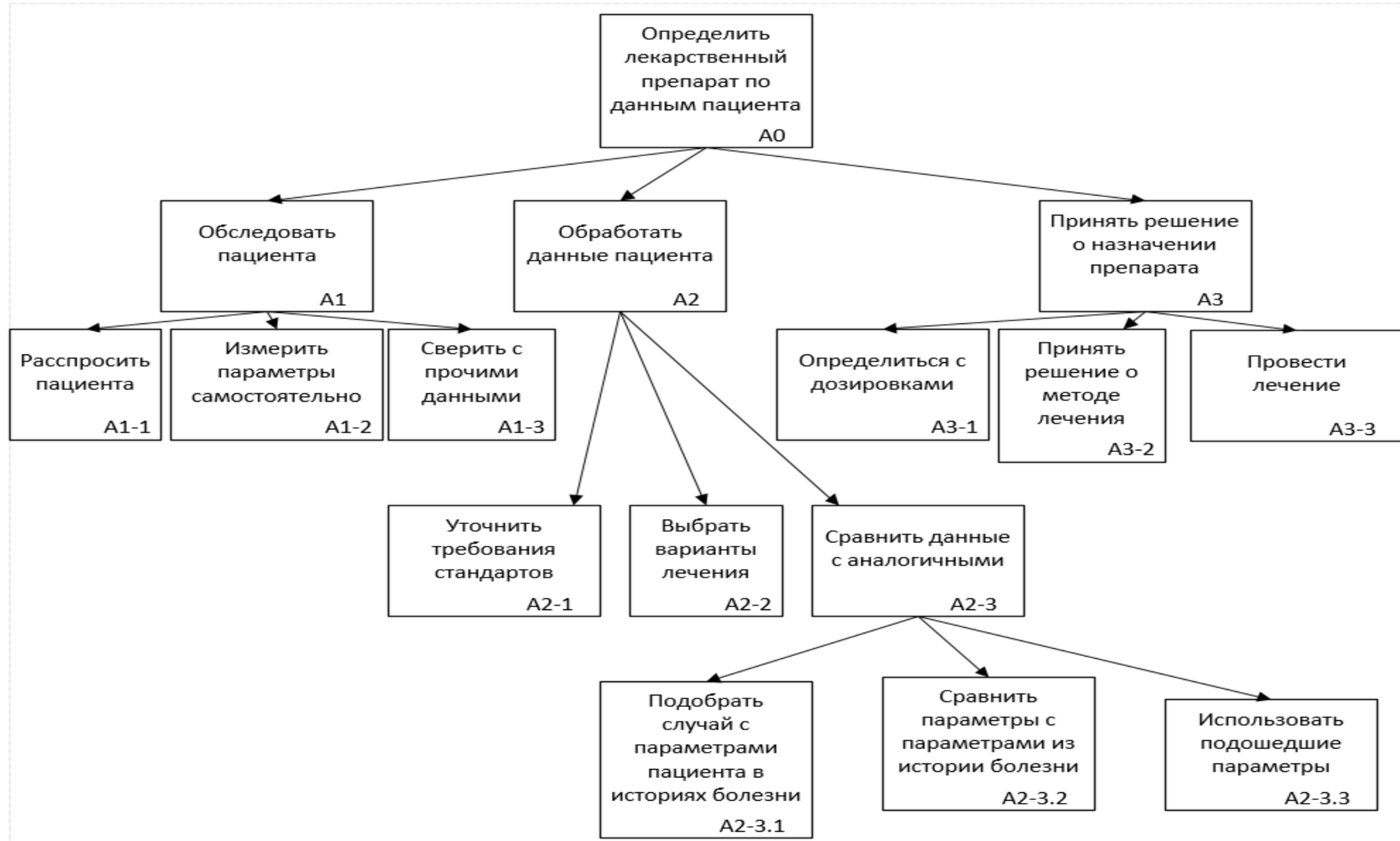


Схема одной итерации

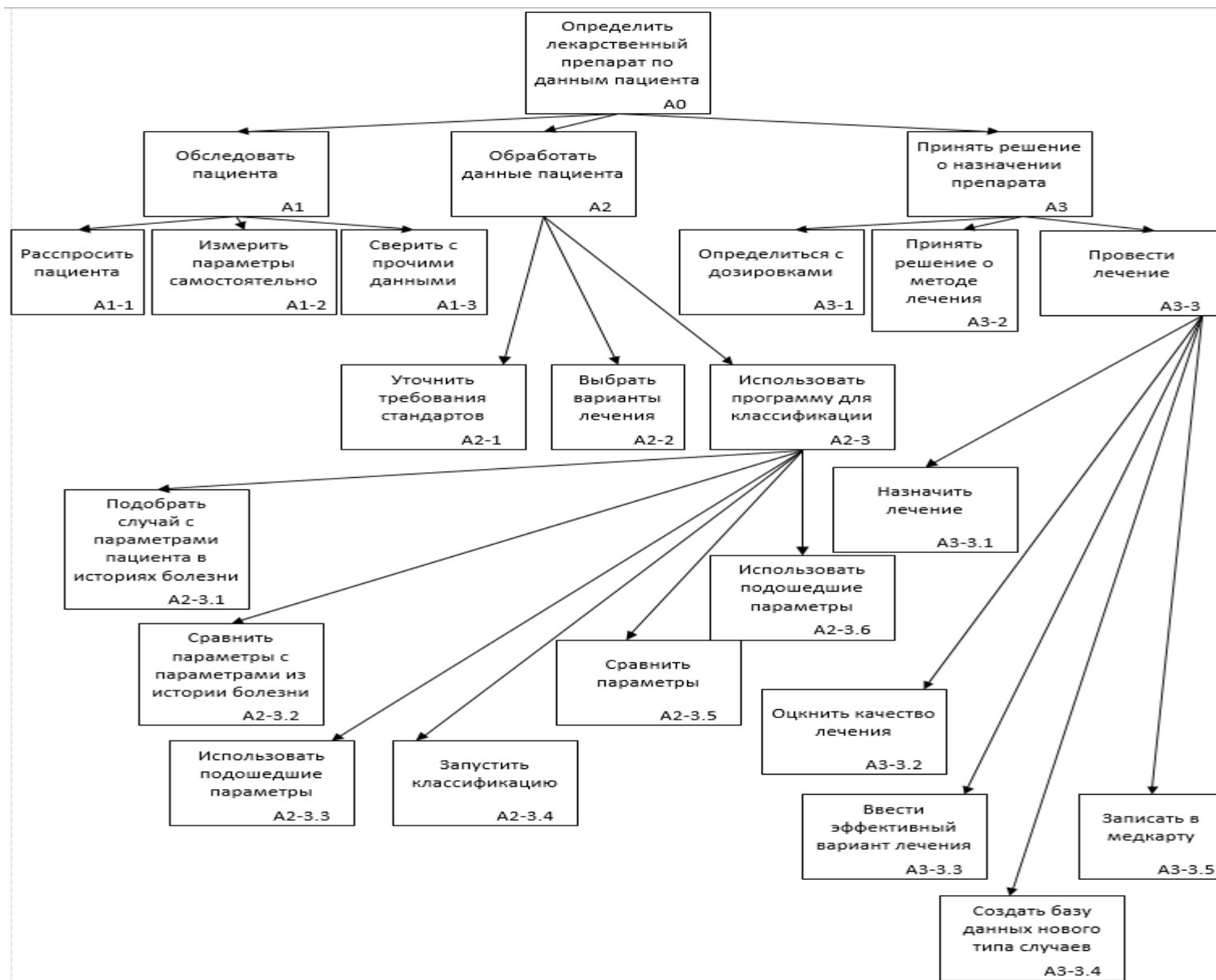
1.2 Таблица требований

Пользовательское требование	Функциональное требование	Программный компонент	Номер итерации
Хранение параметров пациентов	Таблица параметров пациентов	База данных SQLite с таблицей данных	1
Обучение классификатора на основе метода распознавания образов при помощи обучающей выборки.	Алгоритм расчета параметров (весов) классификатора	Программный код на ЯП2 Python, выполняющий расчет параметров (весов) классификатора	2
Создание выборки для новой медицинской задачи с произвольными параметрами и препаратами	Создание таблицы с произвольными атрибутами параметров списка лекарств	Окно «Создать новую базу данных» одноименное окно	3
Классификация записей пациентов их таблицы с целью выяснения подходящего для лечения препарата	Возможность применения классификатора к записям	Режим главного окна для классификации записей таблицы классифицируемой БД и пункт меню для его запуска	3

1.3 Карта бизнес-процессов в AS-IS (как есть)



1.4 Карта бизнес-процессов в ТО-ВЕ (как будет)



2.1 Разработка программы: 1-ая итерация – проектирование данных

номер_записи	препарат	пациент	температура	уровень_боли	показатель_Чайльда_Пью
1	Aspirin	0	36.3	94.0	2.0
2	Aspirin	1	36.4	93.0	2.1
3	Aspirin	2	36.2	93.0	2.1
4	Aspirin	3	36.6	93.0	2.1
5	Aspirin	4	36.0	94.0	2.0
6	Aspirin	5	36.4	91.0	2.0
7	Aspirin	6	36.2	91.0	2.0
8	Aspirin	7	36.5	94.0	2.0
9	Aspirin	8	36.6	96.0	2.1
10	Aspirin	9	36.3	94.0	2.3

Таблица данных

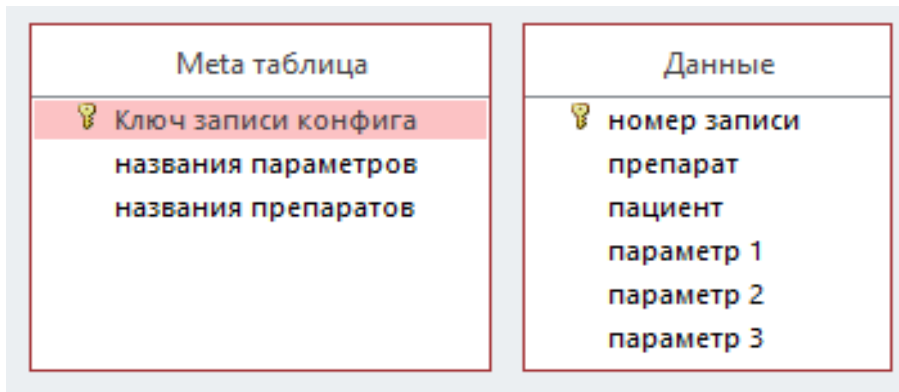


Схема базы данных

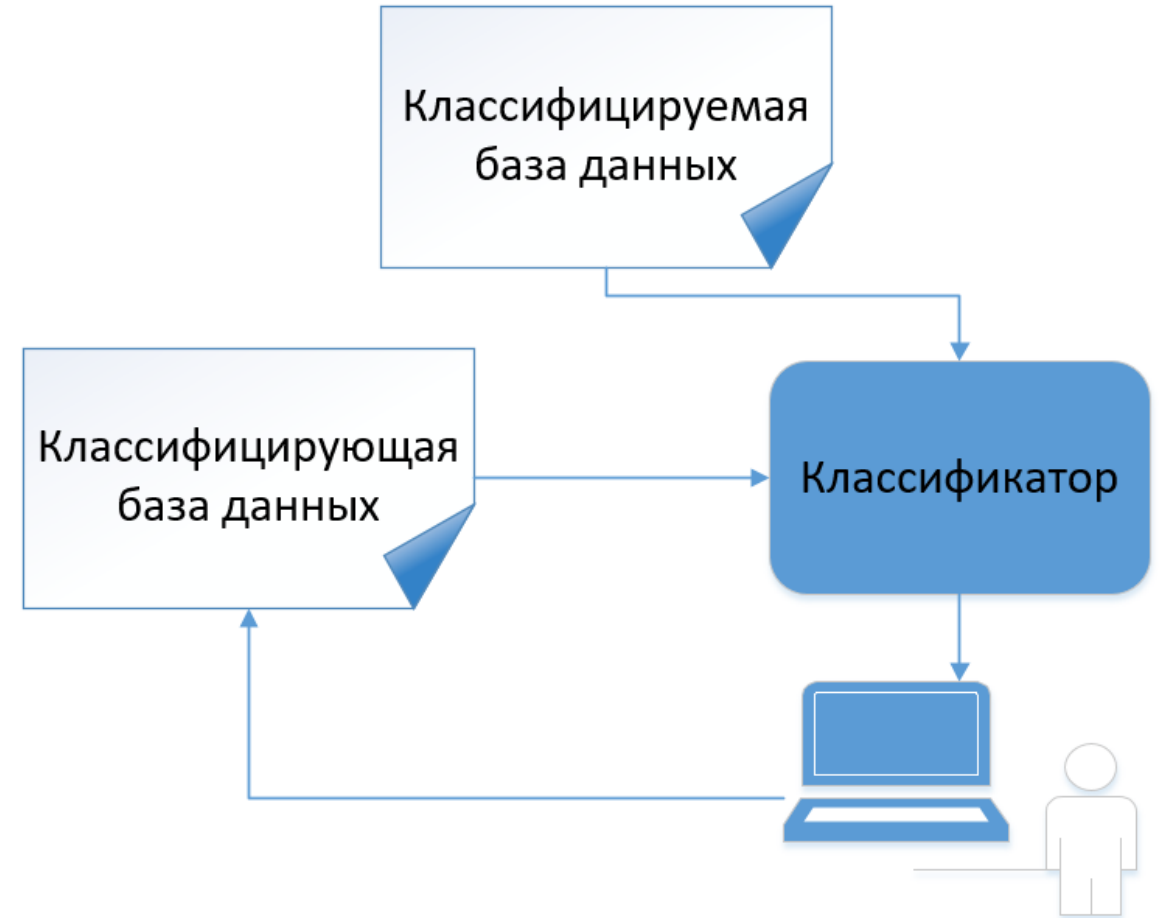
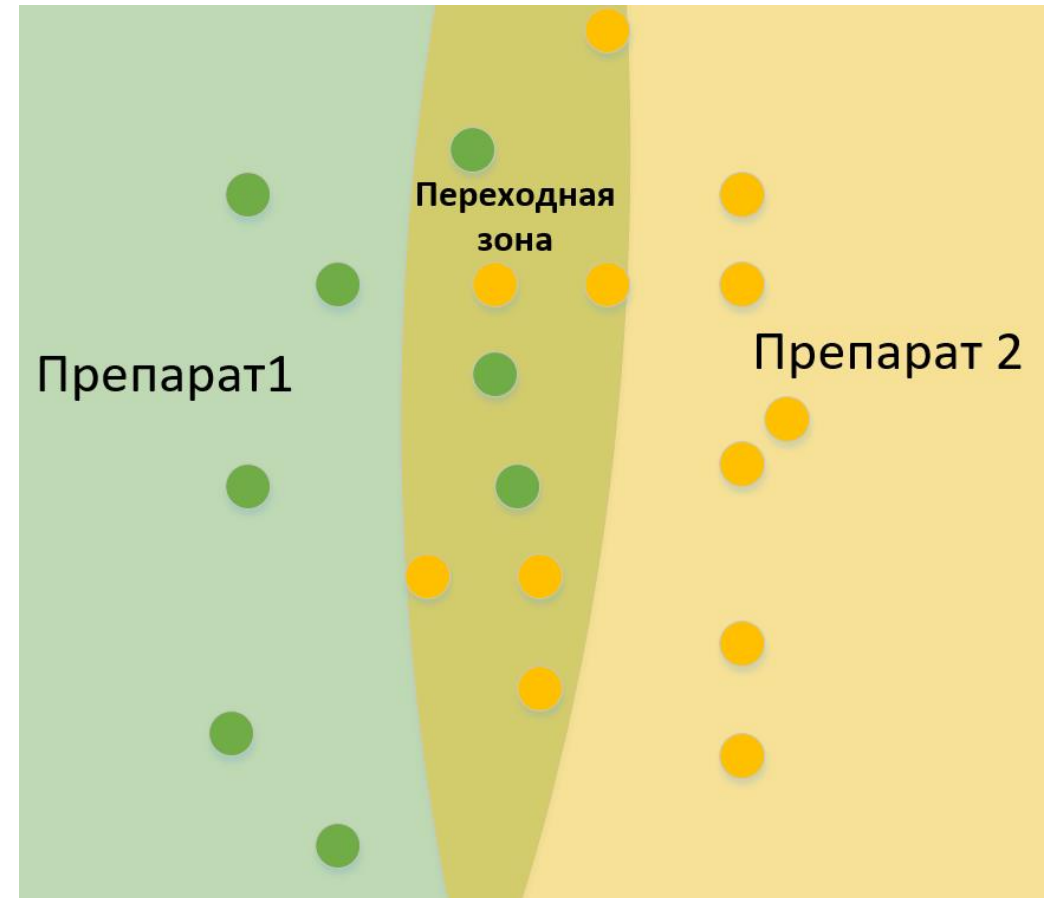


Схема использования данных

2.2 Разработка программы: 2-ая итерация – использование математических методов



Пространство параметров пациентов



Пересечение зон препаратов

2.3 Разработка программы: 2-ая итерация – использование математических методов

Найти веса классификатора по классифицирующей выборке

Рассчитать расстояния по каждому параметру



Найти ошибки по каждому параметру как процент ближайших элементов чужого класса (другого препарата)



Вычислить веса для каждого параметра

Применить классификатор и пополнить классифицирующую выборку

Найти ближайший элемент классифицирующей выборки с использованием весов



По решению добавить элемент в классифицирующую выборку, вычислив новые веса

2.4 Разработка программы: 2-ая итерация – использование математических методов

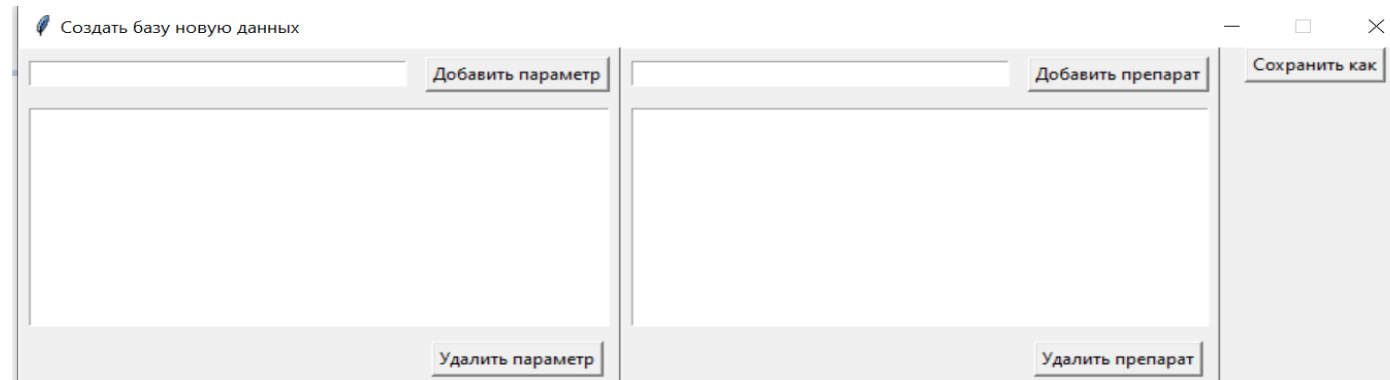
$r = |p_1 - p_2| / \max(p_1, p_2)$ – формула для расчета расстояния (r) выборки по параметру, например, температуре.

$W = (1 - r_i) / \sum_{j=1}^k r_j$ – формула расчета веса i-того параметра из k параметров.

$R = \sum_{i=1}^k W_i r_i$ – формула вычисления расстояния классификатора по k параметрам.

$R = \sum_{i=1}^k \frac{r_1}{W_1} = \min r_3$ – формула принципа минимального расстояния для k параметров.

2.5 Разработка программы: 3-ая итерация – проектирование пользовательского интерфейса: режимы создания и редактирования баз данных

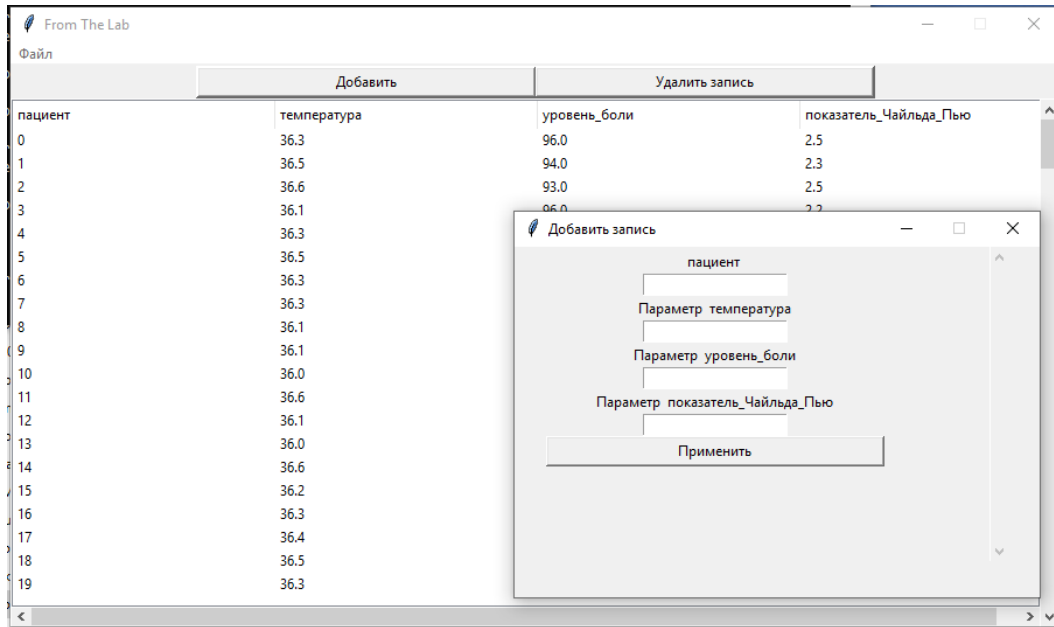


Режим создания базы данных с новыми наборами параметров и применяемых препаратов

номер_записи	препарат	пациент	температура	уровень_боли	показатель_Чайльда_П
1	Aspirin	0	36.2	91.0	2.1
2	Aspirin	1	36.6	94.0	2.1
3	Aspirin	2	36.6	96.0	2.3
4	Aspirin	3	36.1	95.0	2.2

Режим редактирования баз данных

2.6 Разработка программы: 3-ая итерация – проектирование пользовательского интерфейса: режимы ввода и классификации



Режим ввода записей (параметров) с формой

препарат	пациент	температура	уровень_боли	показатель_Чайльда_Пью
Aspirin	0	36.3	96.0	2.5
Aspirin	1	36.5	94.0	2.3
Aspirin	2	36.6	93.0	2.5
Aspirin	3	36.1	96.0	2.2
Aspirin	4	36.3	93.0	2.1
Aspirin	5	36.5	95.0	2.2
Aspirin	6	36.3	91.0	2.1
Aspirin	7	36.3	90.0	2.5
Aspirin	8	36.1	91.0	2.5
Aspirin	9	36.1	95.0	2.5
Aspirin	10	36.0	93.0	2.5
Aspirin	11	36.6	95.0	2.2
Aspirin	12	36.1	91.0	2.2
Aspirin	13	36.0	92.0	2.2
Aspirin	14	36.6	96.0	2.2
Aspirin	15	36.2	90.0	2.2
Aspirin	16	36.3	96.0	2.4
Aspirin	17	36.4	90.0	2.1
Aspirin	18	36.5	93.0	2.5
Aspirin	19	36.3	90.0	2.0

Режим классификации.

Заключение

Проведен анализ требований по созданию медицинской информационной системы, предлагающей методику лечения на основе входных параметров пациента.

Проделано моделирование процессов системы в нотации IDEF0 на верхнем уровне с нулевого до первого уровня декомпозиции. На нижнем уровне моделирование процессов велось с применением нотации IDEF3. Построены карты процессов.

Создан и протестирован прототип программы по итерационной модели. Проведен тест классификации исследуемой выборки улучшенным методом ближайшего соседа на основании значений параметров из классифицирующей выборки и весов, вычисленных на основе этих параметров, автоматическое сопоставление идентичных параметров двух схожих с нахождением идентичных элементов.

Спасибо за внимание