

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ**  
**ФГБОУ ВО "КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ**  
**ИМЕНИ И.Т. ТРУБИЛИНА"**

---

**Факультет прикладной информатики**  
Кафедра компьютерных технологий и систем  
Кострома Дмитрий Сергеевич

ДОПУСКАЕТСЯ К ЗАЩИТЕ  
**Руководитель образовательной программы**  
д.т.н., профессор В.И. Лойко

\_\_\_\_\_  
(подпись)  
« \_\_ » \_\_\_\_\_ 2019 г.

ДОПУСКАЕТСЯ К ЗАЩИТЕ  
**Заведующий кафедрой КТС**  
д.т.н., профессор В.И. Лойко

\_\_\_\_\_  
(подпись)  
« \_\_ » \_\_\_\_\_ 2019 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

На тему: «Исследование и разработка семантических ядер нозологических образов и создание интеллектуального интерактивного справочника болезней»

по направлению: 09.04.02 «Информационные системы и технологии»

Руководитель:

д.э.н., профессор

\_\_\_\_\_  
(подпись)

Луценко Е.В.

Нормоконтроллер:

д.э.н., профессор

\_\_\_\_\_  
(подпись)

Луценко Е.В.

Краснодар  
КубГАУ  
2019

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФГБОУ ВО "КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ  
ИМЕНИ И.Т. ТРУБИЛИНА"

---

**Факультет прикладной информатики**  
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ  
Заведующий кафедрой КТС  
д.т.н, профессор  
В.И. Лойко \_\_\_\_\_  
(подпись)  
« \_\_ » \_\_\_\_\_ 2019 г.

**ЗАДАНИЕ НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ**  
по направлению 09.04.02 Информационные системы и технологии

студенту учебной группы: ИТ1741  
Костроме Дмитрию Сергеевичу

1. Тема работы

Исследование и разработка семантических ядер нозологических образов и создание интеллектуального интерактивного справочника болезней

2. Цель работы

Разработка инструментария для улучшения работы медицинского учреждения.

3. Исходные данные и методические указания

Методические рекомендации по подготовке к государственной итоговой аттестации: магистров по специальности 09.04.02 «Информационные системы и технологии» / сост. В. И. Лойко [и др.]. – Краснодар: КубГАУ, 2019. – 77 с

4. Содержание итоговой квалификационной работы

- Исследование структуры предприятия
- Моделирование объекта исследования и проведение экспериментов
- Разработка информационной системы
- Оценка эффективности разработанной информационной системы

5. Перечень основной рекомендуемой литературы

- Методические рекомендации по подготовке к государственной итоговой аттестации / сост. В.И. Лойко [и др.]. - Краснодар: КубГАУ, 2019. – 77 с.

Задание рассмотрено на заседании кафедры (протокол № \_\_ от " \_\_ " \_\_\_\_\_ 201\_г.)

Руководитель: д.э.н., профессор Луценко Е.В. \_\_\_\_\_

(подпись)

Задание к исполнению принял: " \_\_ " \_\_\_\_\_ 201\_\_ г.

\_\_\_\_\_  
(подпись студента)

## АННОТАЦИЯ

Выпускная квалификационная работа содержит 80 страниц, 31 рисунок, 7 таблиц, 16 используемых источников.

Ключевые слова: ЭЙДОС, СЕМАНТИЧЕСКОЕ ЯДРО, ДИАГНОЗ, ПАЦИЕНТ, НОЗОЛОГИЧЕСКИЙ ОБРАЗ.

**Цель исследования** – разработка инструментария для улучшения работы медицинского учреждения.

**Объект исследования** – интеллектуальный интерактивный справочник болезней.

**Предмет исследования** – адаптивная интеллектуальная модель семантических ядер нозологических образов.

**Научная новизна исследования:** разработан интерактивный интеллектуальный справочник болезней, отличающиеся функциями от обычных справочников.

**Результаты работы:** в результате проведения исследования было выяснено, что существует большое количество систем по автоматизации. Большинство из них обладают обширным функционалом, избыточным для данного предприятия – объекта автоматизации. Приобретение данных систем не является рациональным. Поэтому было принято решение разработать самостоятельный программный продукт, позволяющий улучшить работу больницы за счет упрощения и ускорения процесса принятия решений.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	5
1. ИССЛЕДОВАНИЕ СТРУКТУРЫ ПРЕДПРИЯТИЯ .....	7
1.1. Краткая характеристика объекта исследования интерактивного интеллектуального справочника болезней .....	7
1.2. Эффективная модель управления качеством медицинских услуг .....	10
1.3. Принципы и модель разрабатываемой системы .....	14
1.4. Постановка задачи .....	17
1.5. Выводы .....	18
2. МОДЕЛИРОВАНИЕ ОБЪЕКТА ИССЛЕДОВАНИЯ И ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТОВ .....	22
2.1. Обоснование выбора метода решения проблемы .....	22
2.2. Суть математической модели и частные критерии .....	25
2.3. Интегральные критерии и принятие решений .....	31
2.4. Выводы .....	33
3. РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ .....	34
3.1. Подготовка исходных данных .....	34
3.2. Разработка системно-когнитивных моделей и численное решение задач .....	37
3.3. Описание алгоритма работы системы .....	64
3.4. Выводы .....	66
4. ОЦЕНКА ЭФФЕКТИВНОСТИ РАЗРАБОТАНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ .....	67
4.1. План внедрения .....	67
4.2. Оценка эффективности .....	68
4.3. Определение плановой себестоимости проведения работ .....	69
4.4. Вывод .....	71
ЗАКЛЮЧЕНИЕ .....	72
СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ .....	73
ПРИЛОЖЕНИЕ 1. ЛИСТИНГ ПРОГРАММЫ .....	75

## ВВЕДЕНИЕ

На предприятиях сегодня активно развиваются информационные технологии, внедряется программное обеспечение. Это информационные порталы, компьютерные сети, системы электронного документооборота, общие базы данных, информационные системы для коммунальных хозяйств и геоинформационные системы.

Разработка и внедрение ИТ-технологий является стратегически важной задачей особенно в сфере медицинских услуг, где компьютеры иногда могут сделать гораздо больше человека.

Одной из ключевых проблем, стоящих перед медициной, это своевременно вынесенный правильный диагноз. За все время существования медицины, человечество накопило немало знаний в этой области. По этим знаниям обучаются новые специалисты. Но информации настолько много, что своевременно найти в ней нужную иной раз не получается, а ведь это может очень дорого стоить человеку, пришедшему на прием к доктору. В этом специалисту приходит на помощь компьютер. Информационные технологии, обучение на информационных базах отлично справляются с задачей идентификации болезни и предоставлении наиболее подходящей информации.

Сегодня предпосылкой для продвижения в области информационных технологий является широкое внедрение стандартов и технологий информационных систем, используемых как для аппаратных, так и для программных продуктов. Следует особо подчеркнуть, что сегодня успешная реализация значимых проектов в области информационных технологий, управления, информации и телекоммуникаций невозможна без согласования разработок с существующими стандартами в области информационных систем и в некоторых случаях разработки новых стандартов.

Эффективность лечения напрямую зависит от процесса установления диагноза посредством ИС, которая должна обеспечивать оперативный сбор и правильную интерпретацию данных из различных информационных систем.

Мед учреждениям нужны удобные, высокоскоростные средства доступа, просмотра и анализа больших объемов информации, которые можно найти в транзакционных системах или объединить в хранилищах данных. Правильно разработанные инструменты анализа данных позволят преобразовать накопленные данные в полезные знания и использовать их в процессе принятия решений.

Медицинские структуры относятся к особому классу систем - адаптивным самоорганизующимся системам, то есть системам, которые автоматически меняют алгоритм своей работы и, при необходимости, структуру, чтобы поддерживать или достигать оптимального состояния при изменении внешних условий. Самоорганизующиеся системы должны быть открытыми и постоянно обмениваться информацией с окружающей средой. Как адаптивная система, предпринимательские структуры самостоятельно выбирают и ставят цель своего существования и развития, они способны изменять структуру, параметры и алгоритм функционирования и управления, чтобы обеспечить достижение поставленной цели, то есть они самостоятельно развивающиеся системы.

## 1. ИССЛЕДОВАНИЕ СТРУКТУРЫ ПРЕДПРИЯТИЯ

### 1.1. Краткая характеристика объекта исследования интерактивного интеллектуального справочника болезней

На основе заложенных данных для обучения интеллектуальной системы, формируются семантические ядра нозологических образов. Опираясь на эти рассчитанные ядра, система способна прогнозировать похожие случаи по введенным критериям. В данном справочнике критериями будут являться классы симптомы, диагностика и лечение. В основном будут использоваться симптомы.

Данная система на основе введенных в нее симптомов по одному пациенту или же по группе пациентов, с помощью загрузки файла с данными, способна прогнозировать наиболее вероятные диагнозы, а вместе с ними и выводить соответствующую диагностику и общее лечение данного недуга.

Интерактивный справочник имеет возможность пополняться, а также совершенствовать свою точность в определении диагноза. Чем больше база данных болезней загружена в справочник, тем точнее его определения.

Данный медицинский справочник призван помочь малоопытным специалистам общего профиля диагностированию. Особенно эффективным справочник будет при столкновении с редкими заболеваниями. Если симптомы пациента будут указывать на редкую болезнь, умный справочник непременно сообщит об этом доктору. При повседневной рутине, даже опытный специалист может не заподозрить редкую болезнь. Разрабатываемый же справочник не только сообщит о вероятном заболевании, но и предоставит информацию относительно диагностировании этого недуга.

Данный справочник можно внедрить практически в любое медицинское учреждение. Дальнейшее его рассмотрение целесообразно проводить на реальном примере.

В качестве медицинского учреждения для внедрения данного справочника в тестовом режиме и его последующего обучения будем рассматривать Краевую клиническую больницу №1 Очаповского. После окончания тестирования системы в этом учреждении, доработанную и более точную версию можно будет внедрить в любое медицинское учреждение.

Для разработки данной системы будем рассматривать структуру ККБ№1 так как она имеет схожую структуру медицинской части с другими аналогичными учреждениями.

Всю свою 200-летнюю историю Краевая клиническая больница №1 всегда была лидером среди ведущих учреждений страны по оказанию высокотехнологичной медицинской помощи с применением передовых инновационных технологий диагностики и лечения.

В сфере высокотехнологичного лечения заболеваний сердечно-сосудистой системы Первая краевая больница, по данным академика РАН, главного кардиохирурга России Лео Бокерия, всегда занимала почетные первые места среди медицинских центров Москвы, Санкт Петербурга и Новосибирска. А по мнению заслуженного врача РФ и главного торакального хирурга российского Минздрава Петра Яблонского, клиника – еще и лидер в стране по количеству легочных операций.

Специализированная медицинская помощь, в том числе высокотехнологичная, в ГБУЗ НИИ-ККБ №1 оказывается амбулаторно по 52 профилям, стационарно по 35 профилям. Некоторые виды специализированной медицинской помощи, оказываемые в ГБУЗ ККБ №1, не имеют аналогов в крае. Это кардиохирургия, торакальная хирургия, сосудистая хирургия, вертебрология, комбустиология, генетика, нефрология, детская кардио- и сосудистая хирургия, детская комбустиология.

Сотрудники Краевой клиники удостоились 4-х Премий «Призвание» в различных номинациях. В многолетней истории престижной награды – это уникальный случай, что свидетельствует о высочайшем уровне кубанского здравоохранения в целом.

Номинация «За проведение уникальной операции, спасшей жизнь человека» Лауреат– доктор Владимир Алексеевич Порханов. Награжден за проведение двух уникальных для российской и мировой практики операций по лечению тяжелых поражений легких и трахеи.

Номинация «За проведение уникальной операции, спасшей жизнь человека». Лауреаты – группа врачей под руководством доктора Вадима Николаевича Бодни. Награждены за спасение пациента, насквозь пронзенного арматурой.

Номинация «За проведение уникальной операции, спасшей жизнь человека» Лауреаты – группа хирургов под руководством доктора Кирилла Олеговича Барбухатти. Награждены за проведение операции по спасению пациента с множественными колото-резаными ранами.

Номинация «За создание нового метода лечения» Лауреаты – группа врачей-хирургов под руководством доктора Сергея Борисовича Богданова. Награждены за создание технологии одномоментной послойной пластики лица единым лоскутом кожи пациентам с тяжелыми ожогами лица.

На сегодняшний день IT-технологии применяются во всех областях медицины и здравоохранения. Несмотря на невысокий уровень развития информационных систем в среднем по отрасли, общество морально готово к внедрению новых технологий. Прогресс в информатизации здравоохранения приводит к тому, что медицинские знания становятся всеобщим достоянием. В настоящий момент информационные системы и технологии просто необходимы здравоохранению. Несомненно, внедрение IT-технологий в медицине положительно сказывается на работе здравоохранения в целом. Для пациентов это улучшение качества оказываемых услуг за счет информационной

связи врачей и специалистов, которые участвуют в лечении, также оперативная передача информации между различными учреждениями. Пациенты получают медицинские записи в печатном виде, что является несомненным для них плюсом, а специалисты перераспределяют силы в повседневной работе, сокращая время «бумажной работы». Внедрение системы способствует росту уровня сервиса, также возрастает доступный объем справочной информации для пациентов. После внедрения единой базы данных можно быть уверенным в сохранности медицинской карты пациента и истории болезни. В отличие от бумажного аналога, электронную медицинскую карту пациент не унесет с собой и не потеряет. А государственные стандарты в сфере IT-технологий дают гарантии безопасности персональных данных. Все вышеперечисленные преимущества направлены в первую очередь на точность, качество, оперативность и слаженность работы системы, а также на сокращение ошибок, времени и затрат человеческих ресурсов.

Информационные системы и технологии – это полезный инструмент для повышения качества и эффективности медицинской помощи. Но стоит заметить, что их применение требует детального подхода к подготовке медицинского персонала, организации структуры непосредственно медицинской помощи и управлением ею. Нерешенные проблемы внедрения IT технологий в медицину требуют дополнительного изучения.

## 1.2. Эффективная модель управления качеством медицинских услуг

Качество медицинской деятельности, как профессиональной составляющей медицинской помощи, должно объективно отражать технические, организационно-методические и социальные аспекты. Данное понятие определено как совокупность компонентов системы профессионального оказания медицинской помощи (технический, организационный, методический и социальный), способствующих формированию у пациентов чувства удовлетво-

ренности в ходе взаимодействия с медицинскими работниками, при достаточном ресурсном, квалификационном и технологическом обеспечении.

По оценкам специалистов, в профессиональном сообществе основным инструментом контроля является оценка, направленная на выявление нарушений и дефектов в лечебно-диагностическом процессе, и применение административных мер. Таким образом, оценка качества медицинской деятельности – это элемент и инструмент контроля.

Сложившаяся в нашей стране практика показывает, что как бы внешне ни различались подходы и методы оценки качества медицинских услуг, их объединяет один и тот же принцип: изучение лечебно-диагностического процесса или отдельных его этапов, которые сопоставляются с какими-то формализованными или неформализованными критериями. В основе всех предлагаемых методик лежит, как правило, метод экспертных оценок, в остальном же они отличаются лишь набором показателей и способами их расчетов.

Унифицировать подходы к оценке качества медицинских услуг невозможно без учета современных требований к оказанию профессиональной медицинской помощи: соответствие медицинской деятельности требованиям потребителя, развитие внутренней, внешней и независимой оценки как новых элементов системы гарантии медицинских услуг, становление механизмов общественно-государственного управления, создание инструментов профессионально-общественной аккредитации и лицензирования профессиональной медицинской деятельности и т.д.

Системный подход к оценке качества требует максимально возможного учета всех аспектов проблемы в их взаимосвязи и целостности, выделения главного и существенного, определения характера связей между компонентами. При системном подходе процесс оказания медицинской помощи и его оценка рассматривается как открытая динамическая система, состоящая из взаимосвязанных подсистем, оказывающих влияние на вышестоящую систему, и сама подверженная ее влиянию. В этом плане процесс оказания меди-

цинской помощи состоит из взаимодействия внешних систем и внутренних подсистем. Следовательно, ее результат определяется множеством внешних и внутренних взаимосвязанных процессов.

Построение модели, объективно отображающей взаимоотношения элементов и характеристик медицинской деятельности, является приоритетной задачей в решении сопряженных проблем, возникающих на пути научных исследований и практической деятельности.

Исходя из определения как совокупности компонентов системы профессионального оказания медицинской помощи (технический, организационный, методический и социальный), способствующих формированию у пациентов чувства удовлетворенности в ходе взаимодействия с медицинскими работниками, при достаточном ресурсном, квалификационном и технологическом обеспечении, разработана модель системно-ориентированного подхода к оценке качества медицинских услуг.

В контексте системно-ориентированного подхода к оценке качества медицинских услуг организация здравоохранения представляется в виде уровневой производственной системы, содержащей совокупность взаимосвязанных между собой компонентов, в которой в ходе управленческого цикла (планирования, организации, руководства и контроля) ресурсы организации преобразуются в определенный результат деятельности на выходе, заданный поставленными целями. Для данных организаций они могут быть определены как: 1) обеспечение заданного стандартами качества предоставляемых медицинских услуг; 2) достижение ключевых показателей эффективности работы медицинской организации.

В наиболее востребованной модели выделяются 3 условных уровня: подсистема лицензирования и аккредитации; технологическая и экономическая подсистема; деонтологическая подсистема. Подсистема лицензирования и аккредитации определяет оценку качества медицинских услуг по материальной (технический компонент) и кадровой (организационный компонент)

структуре ее реализации; технологическая - по степени соответствия технологий медицинского труда определенным стандартам; экономическая - по стоимостным характеристикам проведенных медицинских процедур; деонтологическая - по степени удовлетворенности пациентов и, что немаловажно, удовлетворенности врача своей деятельностью. В рамках последней подсистемы описываются морально-этические характеристики и требования к врачу.

На сегодняшний день IT-технологии применяются во всех областях медицины и здравоохранения. Несмотря на невысокий уровень развития информационных систем в среднем по отрасли, общество морально готово к внедрению новых технологий. Прогресс в информатизации здравоохранения приводит к тому, что медицинские знания становятся всеобщим достоянием. В настоящий момент информационные системы и технологии просто необходимы здравоохранению. Несомненно, внедрение IT-технологий в медицине положительно сказывается на работе здравоохранения в целом. Для руководства это постоянный контроль над ходом лечения пациентов, над своими финансами, ведением документации, качеством оказываемой помощи, состоянием коечного фонда и состоянием здоровья пациентов. Персоналу данное внедрение дает возможность участвовать в лечении пациента, находясь в другом городе или стране, обращаться к информационной поддержке по схемам лечения, маршрутизировать заявки на диагностические исследования и врачебные назначения. Для пациентов это улучшение качества оказываемых услуг за счет информационной связи врачей и специалистов, которые участвуют в лечении, также оперативная передача информации между различными учреждениями. Врач может самостоятельно прямо на приеме записывать пациента к другому специалисту. Пациенты получают медицинские записи в печатном виде, что является несомненным для них плюсом, а специалисты перераспределяют силы в повседневной работе, сокращая время «бумажной работы». Внедрение системы способствует росту уровня сервиса, также воз-

растает доступный объем справочной информации для пациентов. После внедрения единой базы данных можно быть уверенным в сохранности медицинской карты пациента и истории болезни. В отличие от бумажного аналога, электронную медицинскую карту пациент не унесет с собой и не потеряет. А государственные стандарты в сфере IT-технологий дают гарантии безопасности персональных данных. Все вышеперечисленные преимущества направлены в первую очередь на точность, качество, оперативность и слаженность работы системы, а также на сокращение ошибок, времени и затрат человеческих ресурсов. Информационные системы и технологии – это полезный инструмент для повышения качества и эффективности медицинской помощи. Но стоит заметить, что их применение требует детального подхода к подготовке медицинского персонала, организации структуры непосредственно медицинской помощи и управлением ею. Нерешенные проблемы внедрения IT технологий в медицину требуют дополнительного изучения.

### 1.3. Принципы и модель разрабатываемой системы

На данном этапе организационная структура медицинского учреждения выглядит следующим образом (см. рис. 1).

Категории медицинских учреждений определены приказом Минздрава России «Об утверждении номенклатуры медицинских организаций». Существуют следующие категории: амбулатория, в том числе врачебная, поликлиники (в том числе детские) государственной и муниципальной систем здравоохранения: консультативно-диагностическая, в том числе детская, медицинской реабилитации, психотерапевтическая, стоматологическая, физиотерапевтическая.

Медицинские организации, в которых располагаются структурные подразделения образовательных и научных организаций, на базе которых осуществляется практическая подготовка медицинских работников (клинические базы), включают в свое наименование слово клинический.

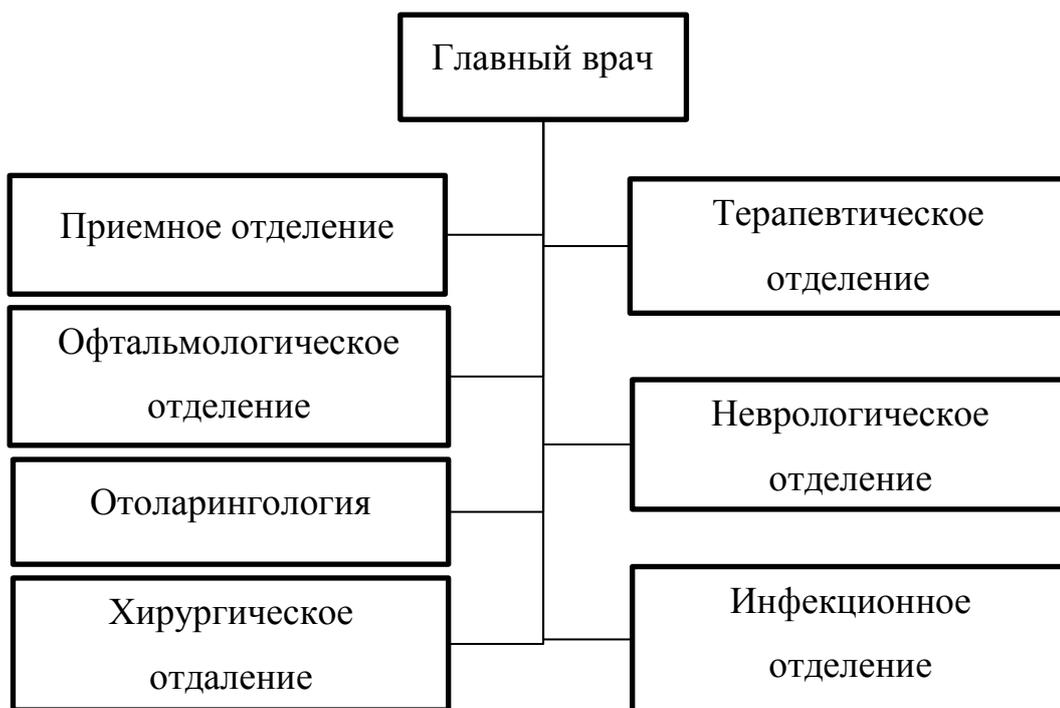


Рисунок 1 – Структура отделений больницы

Виды участков: фельдшерский; терапевтический (в том числе цеховой); врача общей практики (семейного врача); комплексный (участок формируется из населения участка с недостаточной численностью прикрепленного населения (малокомплектный участок) или населения, обслуживаемого врачом-терапевтом врачебной амбулатории, и населения, обслуживаемого фельдшерско-акушерскими пунктами (фельдшерскими здравпунктами); акушерский; приписной.

Фельдшерский участок - 1300 человек взрослого населения в возрасте 18 лет и старше; терапевтический участок - 1700 человек взрослого населе-

ния в возрасте 18 лет и старше (для терапевтического участка, расположенного в сельской местности, - 1300 человек взрослого населения); участок врача общей практики - 1200 человек взрослого населения в возрасте 18 лет и старше; участок семейного врача - 1500 человек взрослого и детского населения; на комплексном участке - 2000 и более человек взрослого и детского населения. Превышение численности прикрепленного населения более чем на 15 процентов от нормативной, не допускается.

Порядок действий, необходимый к реализации в программном продукте:

- пациент приходит к врачу и произносит все свои жалобы;
- на усмотрение специалиста, он уже может воспользоваться системой прогнозирования диагноза, чтобы иметь уже несколько возможных вариантов.

- если по первичным жалобам специалисту не ясна болезнь пациента, он проводит первичный осмотр и дополняет список симптомов пациента;

- специалист заносит все симптомы пациента в таблицу для прогнозирования, и на локальном компьютере решает задачу прогнозирования для этого пациента;

- получив результат программы сообщает пациенту предварительный диагноз. Если специалист согласен с прогнозируемым диагнозом, он может записать его в медицинскую карту пациента и назначить лечение. Если же нет, врач может продолжить осмотр, перенаправить пациента к узконаправленному специалисту либо со всеми собранными данными созвать комиссию компетентных врачей и обсудить диагноз с ними.

Специалист при выборе решения при несогласии с диагнозом, прогнозированным программой, все же может им руководствоваться и, для большей безопасности, провести дополнительное обследование.

Если диагноз вынесен правильно, то информация об этом дополняет имеющуюся модель. Это позволяет отслеживать конъюнктуру и сохранять высокую степень объективности системы.

#### 1.4. Постановка задачи

Необходимо произвести разработку инструментария улучшения работы больницы. Главная задача – определение диагноза пациента исходя из его симптомов. Произвести анализ базы болезней и вычислить потенциальные критерии, присущие тем или иным болезням, которые могут с большой долей вероятности указать на возможную болезнь. Врач должен иметь инструмент быстрого определения возможной болезни пациента. Инструмент, способный в краткие сроки произвести оценку конкретных жалоб пациента и сделать выводы о его возможной болезни. Также система должна быть способна указать наиболее приемлемый способ лечения конкретного пациента. На основе рекомендаций системы, врач будет диагностировать недуг пациента и назначать лечение.

Необходима система, построенная на основе искусственного интеллекта. Собрав информацию из справочников болезней и предоставив ее системе, научим систему определять наиболее достоверные диагнозы и назначать лечение. Данные результаты будем сохранять. Компетентные специалисты будут корректировать работу программы, отмечая, правильные ли вынесен диагноз. Базу данных с корректировками специалистов предоставим программе, тем самым, увеличив достоверность ее решений.

Собрав все реальные данные о жалобах пациентов, система использует их для прогнозирования и принятия решений. Задача решается локально, так как на первом этапе работы программы, нескорректированные прогнозы могут не соответствовать действительности, что может фатально повлиять на пациента. Система выносит решения в разных медицинских областях, что

усложняет их проверку, для чего нередко придется созывать консилиумы врачей, для установления точного диагноза.

## 1.5. Выводы

Многогранная работа не может существовать без объективных данных, полного и всеохватывающего анализа сведений, быстрой и массовой популяризации обнаруженных закономерностей, то есть без налаженной, современной и оперативной ИС.

Преимущества информационной системы очевидны как для всей отрасли здравоохранения в целом, так и для сотрудников. Удобная статистика работы лечебного учреждения, истории болезни, различные планы лечения. В наше время нет необходимости доказывать эффективность информационных систем (ИС), так как они используются более десятка лет и уже доказали свою эффективность не один раз.

Сейчас мировой рынок ИС предлагает много решений. Только в России на данный момент зарегистрировано около 775 различных ИС для здравоохранения. В это число входят самые разнообразные решения: от небольших программ для формирования статистических отчетов до мощных и современных комплексных медицинских информационных систем. Если покупка ИС осуществляется на государственном уровне, то есть на бюджетные средства государства, то руководителям медицинских учреждений следует относиться более ответственно к выбору ИС. При рассмотрении информационных систем, главным фактором, определяющим их преимущество, является объем функционала, который разработчик вложил в ИС. Автоматизация системы здравоохранения, да и систем в целом является затратной. Начальная цена системы оценивается в полмиллиона рублей, и обучение персонала, а также повышение его квалификации увеличивает стоимость внедрения. Но с ус-

вершенствованием ИС и автоматизации бизнес процессов появляется необходимость их применения, так как информационные системы в разы превосходят ручной труд.

При выборе ИС нужно узнать есть ли возможность настроить систему в широком спектре без изменения программного кода. То есть поставщик программного продукта должен показать, как система может адаптироваться к особенностям работы в данном медицинском учреждении без перепрограммирования, а исключительно посредством интерфейса программы или администратора. В любом модуле системы существует потребность в возможности настройки составных частей (для экранных форм, отчетов и иных данных, которые хранятся в базе).

Масштабируемость. Предыдущий критерий тесно связан с таким компонентом как масштабируемость системы. Данное понятие определяет способность системы подстраиваться к определенным потребностям медицинского учреждения по мере ее расширения. Чаще всего под масштабируемостью понимается способность поддерживать количественный рост, будь то увеличение числа пользователей или числа структурных единиц. Поддержка и сопровождение Одним из важных составляющих при отборе ИС является ее техническая поддержка. Вероятнее всего уровень технической поддержки будет зависеть от стоимости системы, но в любом случае данный момент должен быть тщательно рассмотрен. Оптимальное решение вопроса сервисной поддержки можно добиться путем комплексного подхода с применением как экономических, так и организационных мер.

Одним из известных подходов является разделение функций сервисной поддержки между поставщиком и внутренней ИТ-службой лечебного учреждения. Но чаще данное решение доступно в основном для крупных медицинских учреждений. Конечно же, руководитель любой организации дает оценку стоимости предлагаемых программных продуктов. Иногда возникает сложность в сопоставлении стоимости программных решений из-

за различных методов установления цены. Чаще всего стоимостным показателем является одно рабочее место.

Разработка системы целесообразна и актуальна. В современных условиях компьютер, при должном использовании, способен выдавать результаты намного точнее человека.

- специалисту общего назначения (терапевту) необходимо проверить огромное количество информации, чтобы определить область заболевания и направить пациента к специалисту более узкого профиля;

- нередко происходят случаи, когда больного направляют не к тому специалисту. Тот вынужден потратить время, выслушать больного, произвести некоторые проверки, чтобы понять, что недуг не относится к его области работы и отправить пациента обратно к общему специалисту;

- больной может попасть в ситуацию, в которой врачи будут разбирать его недомогание не комплексно, а каждый в своей области знаний, специфики. Таким образом каждый специалист может назначить со своей стороны лечение по своему профилю, что суммарно окажется излишним для пациента. Например пациент будет плохо переносить избыточную долю лекарств. Либо может наблюдаться противоположные случаи, где каждый врач посчитает недуг пациента в области знаний данного врача незначительным и не станет выписывать лекарства;

- без предварительного диагноза пациенту могут быть назначены дополнительные обследования, количество которых с легкостью может уменьшить интеллектуальная система путем повышения их точности;

- специалист несёт на плечах функцию анализа и принятия решений, основываясь лишь на опыте в его медицинской сфере. Опытный врач не сможет передать в полной мере свои знания вновь пришедшему специалисту, вследствие чего потеря ценного кадра может привести к коллапсу системы, завязанной исключительно на человеческие качества конкретного специалиста. Смена сотрудника прерывает процесс накопления опыта.

Использование информационной системы ускорит вынесение диагнозов, упростит назначение лечения, сократит количество ненужных исследований. Повысится точность перенаправления пациентов к узкоспециализированным врачам, что не будет отвлекать от их непосредственной работы. У таких специалистов будет больше свободного времени, в которое они смогут проводить исследования новых болезней, тем самым пополняя базу для системы.

Такие проблемы как следующие будут решаться быстрее. Специалисту общего назначения (терапевту) необходимо проверить огромное количество информации, чтобы определить область заболевания и направить пациента к специалисту более узкого профиля;

Нередко происходят случаи, когда больного направляют не к тому специалисту. Тот вынужден потратить время, выслушать больного, произвести некоторые проверки, чтобы понять, что недуг не относится к его области работы и отправить пациента обратно к общему специалисту;

## 2. МОДЕЛИРОВАНИЕ ОБЪЕКТА ИССЛЕДОВАНИЯ И ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТОВ

### 2.1. Обоснование выбора метода решения проблемы

Автоматизированный системно-когнитивный анализ (АСК-анализ) разработан проф. Е.В. Луценко в 2002 году для решения широкого класса задач идентификации, прогнозирования, классификации, диагностики, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели. АСК-анализ имеет программный инструментальный – универсальную когнитивную аналитическую систему «Эйдос».

Существует много систем искусственного интеллекта. Универсальная когнитивная аналитическая система «Эйдос-Х++» отличается от них следующими параметрами:

- разработана в универсальной постановке, не зависящей от предметной области. Поэтому она является универсальной и может быть применена во многих предметных областях;

- находится в полном открытом бесплатном доступе, причем с актуальными исходными текстами;

- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т.е. она не требует от пользователя специальной подготовки в области технологий искусственного интеллекта;

- обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е.

не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

- содержит большое количество локальных (поставляемых с установкой) и облачных учебных и научных приложений;

- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решения этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний.

### О соотношении содержания понятий: «Данные», «Информация» и «Знания»

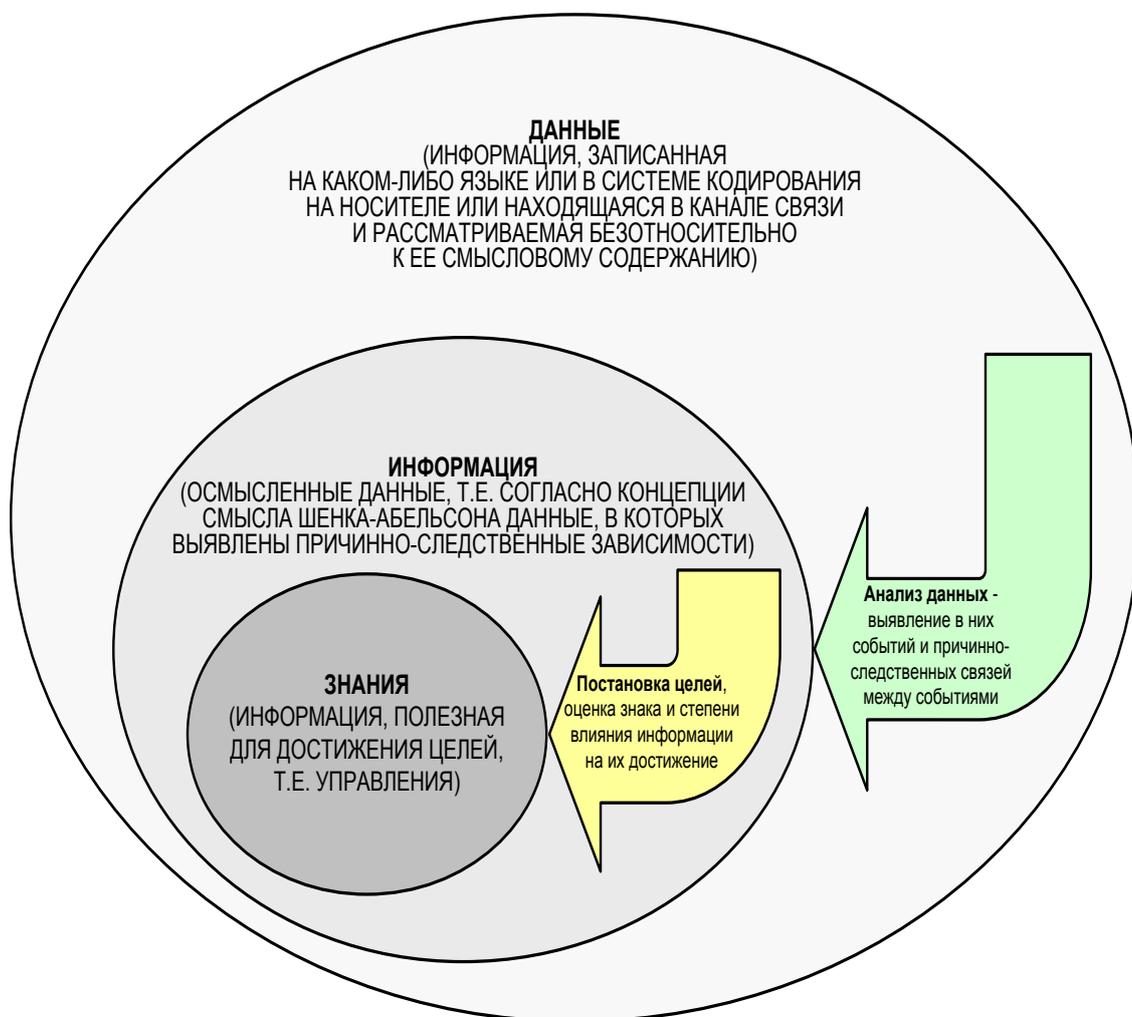


Рисунок 2. О соотношении содержания понятий: «данные», «информация» и «знания» в АСК-анализе

Поддерживает online среду накопления знаний и широко используется во всем мире;

- обеспечивает мульти язычную поддержку интерфейса на 44 языках. Языковые базы входят в инсталляцию и могут пополняться в автоматическом режиме;

- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач классификации, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах

Суть метода АСК-анализа состоит в последовательном повышении степени формализации модели и преобразовании данных в информацию, а ее в знания и решении на основе этих знаний задач идентификации (распознавания, классификации и прогнозирования), поддержки принятия решений и исследования моделируемой предметной области.

На основе, представленных в данной работе (отличающихся по частым критериям), системно-когнитивных моделей, задачи идентификации (распознавание, классификация, прогнозирование, диагностика), проблема изучения и принятия решений моделируемой предметной путем изучения ее системы когнитивная модель исследуется.

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимосвязанных данных, представленных в различных типах шкал (номинальных, порядковых и числовых) и различных единицах измерения.

## Последовательность обработки данных, информации и знаний в системе «Эйдос-X++»

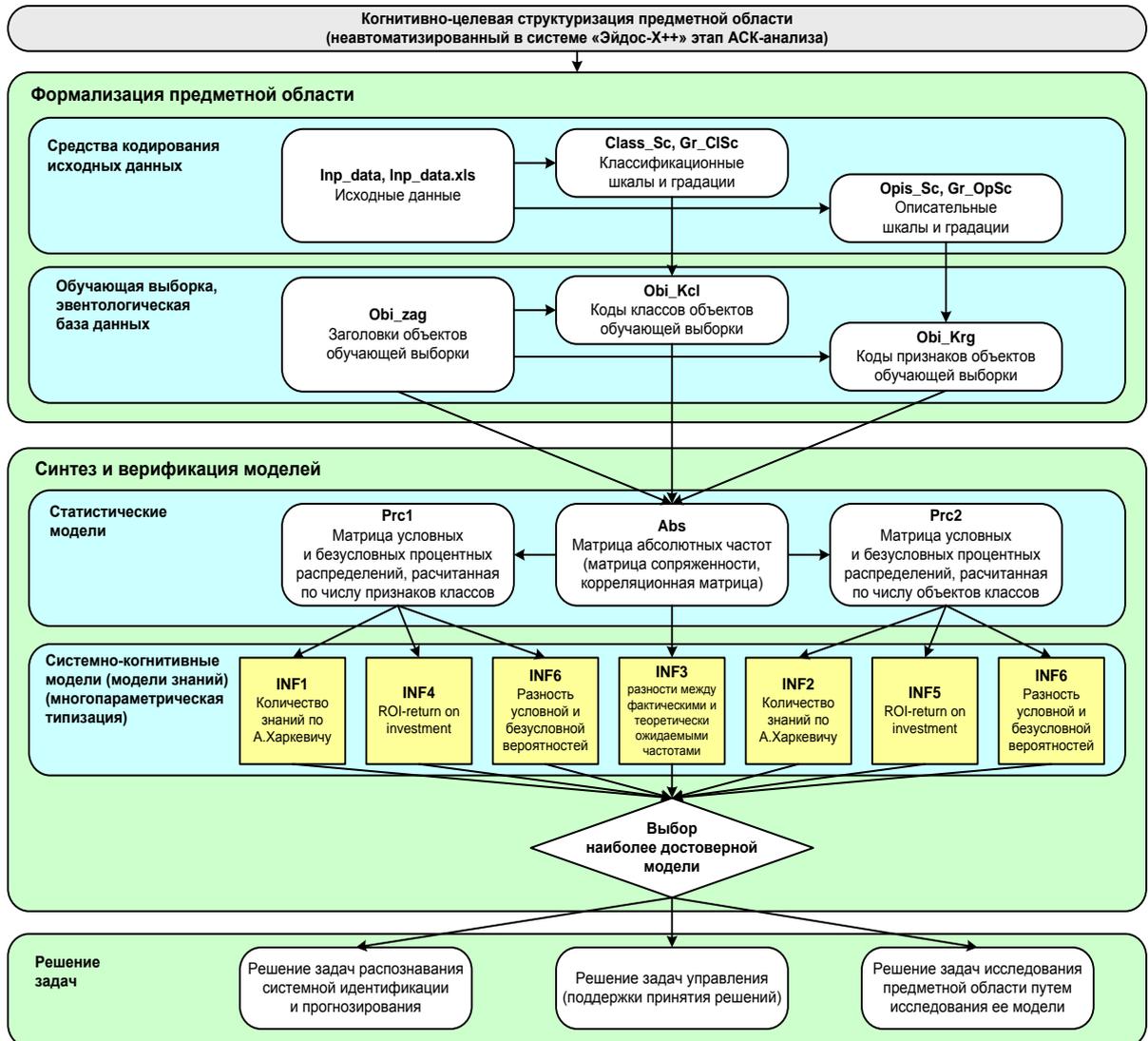


Рисунок 3. Последовательность преобразования данных в информацию, а ее в знания и решения задач в АСК-анализе и системе «Эйдос»

### 2.2. Суть математической модели и частные критерии

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике, как уже говорилось ранее в этой работе, и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых данных, представлен-

ных в различных типах шкал (номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот (таблица 1).

Таблица 1 – Матрица абсолютных частот

		Классы					Сумма
		<i>1</i>	...	<i>j</i>	...	<i>W</i>	
Значения факторов	<i>1</i>	$N_{11}$		$N_{1j}$		$N_{1W}$	
	...						
	<i>i</i>	$N_{i1}$		$N_{ij}$		$N_{iW}$	$N_{i\Sigma} = \sum_{j=1}^W N_{ij}$
	...						
	<i>M</i>	$N_{M1}$		$N_{Mj}$		$N_{MW}$	
Суммарное количество Признаков по классу				$N_{\Sigma j} = \sum_{i=1}^M N_{ij}$			$N_{\Sigma\Sigma} = \sum_{i=1}^W \sum_{j=1}^M N_{ij}$
Суммарное количество объектов обучающей выборки по классу				$N_{\Sigma j}$			$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

На ее основе рассчитываются матрицы условных и безусловных процентных распределений (таблица 2).

Таблица 2 – Матрица условных и безусловных процентных распределений

		Классы					Безусловная вероятность признака
		<i>1</i>	...	<i>j</i>	...	<i>w</i>	
Значения факторов	<i>1</i>	$P_{11}$		$P_{1j}$		$P_{1w}$	
	...						
	<i>i</i>	$P_{i1}$		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		$P_{iw}$	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	<i>M</i>	$P_{M1}$		$P_{Mj}$		$P_{Mw}$	
Безусловная вероятность класса				$P_{\Sigma j}$			

Отметим, что в АСК-анализе и его программном инструментарии интеллектуальной системе «Эйдос» используется два способа расчета матриц условных и безусловных процентных распределений:

1-й способ: в качестве  $N_{\Sigma j}$  используется суммарное количество признаков по классу;

2-й способ: в качестве  $N_{\Sigma j}$  используется суммарное количество объектов обучающей выборки по классу.

Затем на основе таблицы 2 с использованием частных критериев, приведенных таблице 3 рассчитываются матрицы системно-когнитивных моделей (таблица 4).

Таблица 3 – Различные аналитические формы частных критериев знаний

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	через относительные частоты	через абсолютные частоты
<b>ABS</b> , матрица абсолютных частот	---	$N_{ij}$
<b>PRC1</b> , матрица условных и безусловных процентных распределений, в качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу	---	$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$
<b>PRC2</b> , матрица условных и безусловных процентных распределений, в качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу	---	$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$
<b>INF1</b> , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: $N_j$ – суммарное количество признаков по $j$ -му классу. Вероятность того, что если у объекта $j$ -го класса обнаружен признак, то это $i$ -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_i N_j}$
<b>INF2</b> , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: $N_j$ – суммарное количество объектов по $j$ -му классу. Вероятность того, что если предъявлен объект $j$ -го класса, то у него будет обнаружен $i$ -й признак.	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_i N_j}$
<b>INF3</b> , частный критерий: <b>Хи-квадрат</b> : разности между фактическими и теоретически ожидаемыми абсолютными ча-	---	$I_{ij} = N_{ij} - \frac{N_i N_j}{N}$

стотами		
---------	--	--

Таблица 3 – Различные аналитические формы частных критериев знаний  
(продолжение)

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	через относительные частоты	через абсолютные частоты
<b>INF4</b> , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: $N_j$ – суммарное количество признаков по $j$ -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$
<b>INF5</b> , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: $N_j$ – суммарное количество объектов по $j$ -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$
<b>INF6</b> , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: $N_j$ – суммарное количество признаков по $j$ -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$
<b>INF7</b> , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: $N_j$ – суммарное количество объектов по $j$ -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$

Обозначения к таблице 3:

$i$  – значение прошлого параметра;

$j$  – значение будущего параметра;

$N_{ij}$  – количество встреч  $j$ -го значения будущего параметра при  $i$ -м значении прошлого параметра;

$M$  – суммарное число значений всех прошлых параметров;

$W$  – суммарное число значений всех будущих параметров.

$N_i$  – количество встреч  $i$ -м значения прошлого параметра по всей выборке;

$N_j$  – количество встреч  $j$ -го значения будущего параметра по всей выборке;

$N$  – количество встреч  $j$ -го значения будущего параметра при  $i$ -м значении прошлого параметра по всей выборке.

$I_{ij}$  – частный критерий знаний: количество знаний в факте наблюдения  $i$ -го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее  $j$ -му значению будущего параметра;

$\Psi$  – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;

$P_i$  – безусловная относительная частота встречи  $i$ -го значения прошлого параметра в обучающей выборке;

$P_{ij}$  – условная относительная частота встречи  $i$ -го значения прошлого параметра при  $j$ -м значении будущего параметра .

Таблица 4 – Матрица системно-когнитивной модели

		Классы					Значимость фактора
		$1$	...	$j$	...	$W$	
Значения факторов	$1$	$I_{11}$		$I_{1j}$		$I_{1W}$	$\sigma_{1\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{1j} - \bar{I}_1)^2}$
	...						
	$i$	$I_{i1}$		$I_{ij}$		$I_{iW}$	$\sigma_{i\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
	...						
	$M$	$I_{M1}$		$I_{Mj}$		$I_{MW}$	$\sigma_{M\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{Mj} - \bar{I}_M)^2}$
Степень редукции класса		$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Суть этих методов в том, что вычисляется количество информации в факте наличия или определенной степени выраженности того или иного личностного свойства о том, что обладающий им кандидат будет проявлять определенную степень успешности профессиональной деятельности, работая

на той или иной должности. Это позволяет *сопоставимо* и корректно обрабатывать разнородную информацию о респондентах, полученную с помощью различных тестов и других различных источников.

На основе системно-когнитивных моделей, представленных в таблице 4 (отличаются частыми критериями), решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений, а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Для решения этих задач в АСК-анализе и системе «Эйдос» в настоящее время используется два интегральных критерия.

### 2.3. Интегральные критерии и принятие решений

Задача принятия решений представляет собой обратную задачу прогнозирования. Если при прогнозировании на основе значений факторов, воздействующих на объект управления, определяется в какое состояние он под их воздействием перейдет, но при принятии решений наоборот, по желательному (целевому) состоянию объекта управления определяется система значений факторов, обуславливающих переход объекта в это целевое состояние.

Не все модели обеспечивают решение обратной задачи прогнозирования. Для этого они должны обеспечивать многопараметрическую типизацию, т.е. создавать обобщенные образы в будущих состояний объекта управления. Как влияет на поведение объекта управления одно значение фактора отражено в системно-когнитивных моделях. Как влияние система факторов определяется с помощью интегральных критериев. В настоящее время в системе «Эйдос» используется два аддитивных интегральных критерия:

- сумма знаний;
- резонанс знаний.

1-й интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в системе значений факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний:

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид, где  $M$ :

- количество градаций описательных шкал (признаков);
- вектор состояния  $j$ -го класса;
- вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или  $n$ , если он присутствует у объекта с интенсивностью  $n$ , т.е. представлен  $n$  раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» – один раз).

2-й интегральный критерий «Семантический резонанс знаний» представляет собой нормированное суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний и имеет вид где  $M$ :

- количество градаций описательных шкал (признаков);
- средняя информативность по вектору класса;
- среднее по вектору объекта;

- среднеквадратичное отклонение частных критериев знаний вектора класса;
- среднеквадратичное отклонение по вектору распознаваемого объекта.
- вектор состояния  $j$ -го класса;
- вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор). Свое наименование интегральный критерий сходства «Семантический резонанс знаний» получил потому, что по своей математической форме является корреляцией двух векторов: состояния  $j$ -го класса и состояния распознаваемого объекта.

Система «Эйдос» обеспечивает построение информационно-измерительных систем в различных предметных областях. В системе «Эйдос» реализовано большое количество программных интерфейсов, обеспечивающий автоматизированный ввод в систему данных различных типов: текстовых, табличных и графических.

Путем многопараметрической типизации в системе создается системно-когнитивная модель, с применением которой, если модель окажется достаточно достоверной, могут решаться задачи системной идентификации, прогнозирования, классификации, поддержки принятия решений и исследования моделируемого объекта путем исследования его системно-когнитивной модели.

Всем этим и обусловлен выбор АСК-анализа и его программного инструментария интеллектуальной системы «Эйдос» в качестве инструментария решения поставленной проблемы.

## 2.4. Выводы

Интересующая нас система на логическом уровне в полной мере спроектирована. Далее необходимо произвести её реализацию на программном уровне и разработать численные модели.

### 3. РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ

#### 3.1. Подготовка исходных данных

В данном разделе полностью описан численный пример обработки исходных данных, который в настоящее время находится в процессе реализации на сайте. Данный численный пример реализован локально в системе «Эйдос».

Исходные данные размещены на сайте medaboutme.ru – общедоступном справочнике болезней (рисунок 4).

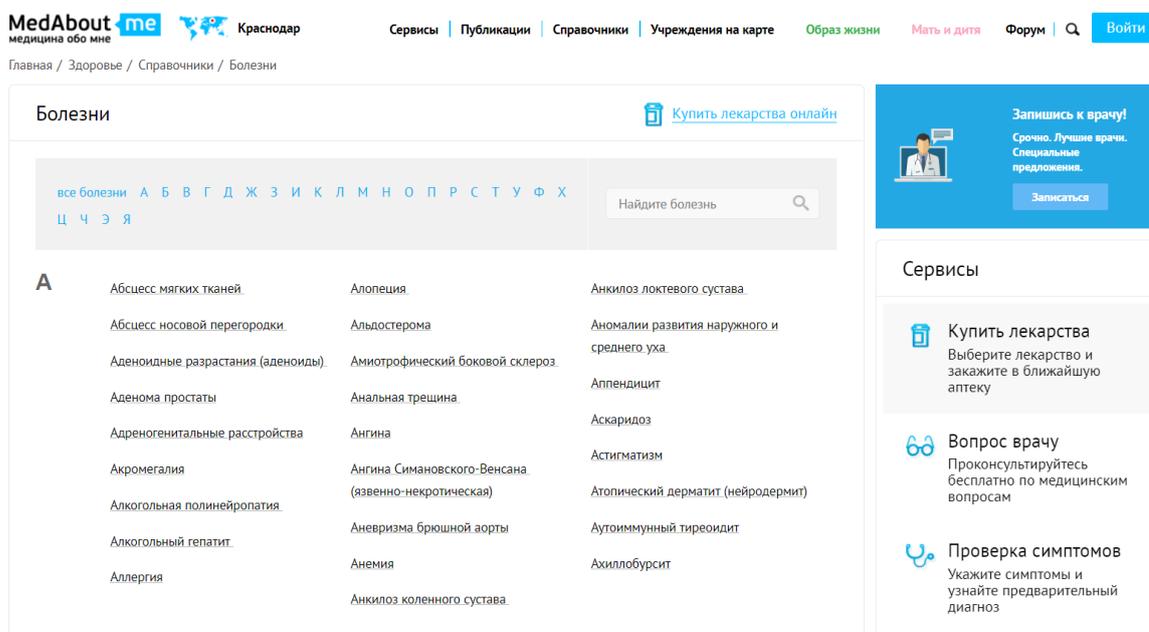


Рисунок 4 – Справочник болезней

Чтобы использовать данные из этого справочника, необходимо разработать одноразовую программу для сбора данных. Ее работа состоит в том, чтобы данные с сайта переписать в формат, удобный для загрузки в систему Эйдос. Для создания такой программы необходимо определить формат загружаемых данных в систему, а также разработать схему этих данных.

Даталогическая модель. Система Эйдос позволяет загружать в нее данные с помощью файлов программы Excel. Для достижения наших целей потребуется всего одна сущность – диагноз (таблица 5).

Таблица 5 – Сущность «Диагноз»

Атрибут	Тип данных	Размер	Описание
Диагноз	Текстовый	50 символов	Название диагноза
Симптомы	Текстовый	Длинный текст	Симптомы присущие данному диагнозу
Диагностика	Текстовый	Длинный текст	Рекомендации по диагностике
Лечение	Текстовый	Длинный текст	Лечение данной болезни

На главной странице сайта клиент оставляет о себе данные, необходимые для связи с ним.

Даталогическая модель данных представляет собой модель данных, которые должны быть описаны и обработаны в области приложения (например, данные производственной зоны, системы учета или всей информации компании) и их отношения друг к другу.

В информатике, особенно в области разработки информационных систем, модели данных и их действия (моделирование данных) служат для поиска и определения структуры данных, подлежащих обработке в системах (в частности, для хранения данных).

Вышеупомянутая модельная градация, основанная на процессе разработки, и в частности обозначения моделей, не применяется на практике. Таким образом, другие термины часто используются в литературе, в публикациях и в обычном использовании; В частности, часто используется только термин «модель данных».

Инфологическая модель. Инфологическая модель - это описание предмета, основанная на анализе семантики объектов и явлений, не ориентиро-

ванная на будущее использование программного обеспечения и технических ИТ-инструментов.

Вне проецируемой базы данных. В соответствии с терминологией в русской литературе внешний дизайн базы данных называется инфологическим дизайном.

Инфологический дизайн - это процесс, который строится путем анализа и идентификации объектов и отношений между ними в предметной области.

Анализ области - это выбор информационных объектов (объектов), настройка необходимых свойств отдельных объектов и выявление связей между ними.

Дизайн интерьера проектировал варианты дизайна, касающиеся большей части запланированной базы данных системы. В соответствии с терминологией, представленной в русской литературе, внутренняя структура базы данных: даталогический дизайн (ориентированный на логическую среду, ориентированный на базу данных дизайн);

Строить инфологическую модель особого смысла нет, так как мы будем работать только с одной таблицей.

Данную программу будем разрабатывать с использованием языка программирования Python (код программы приведен в приложении 1).

Разработанная программа переходит на страницу сайта со справочником, находит на ней все ссылки на болезни (ссылок с болезнями было обнаружено 422), далее переходит по каждой из ссылок на страницы с описанием заболевания. На этих однообразных страницах для каждой болезни, разбирая html страницы, находит соответствующий текст описаний необходимых нам полей. Собрав данные по всем болезням, программа генерирует файл формата xlsx и создает в нем таблицу с данными обо всех найденных болезнях. Фрагмент данных представлен на рисунке 5.

диагноз	симптомы	диагностика	лечение
1. Абсцесс мягких тканей	Абсцесс (гноной) - это ограниченное полостное гнойное воспаление в любых органах и тканях, поэтому он может быть практически любой локализации. Формирование обычно является вторичным, инфекционным, гнойным процессом, возникающим в результате попадания инфекции как извне (например, при инфицировании), так и в результате ее заноса с кровью. Основные жалобы при развитии поверхностного абсцесса: на месте поражения, уплотнение и боль, которая увеличивается по мере его формирования. При благоприятном течении процесс происходит самопроизвольный разрыв капсулы с отхождением гнойного содержимого наружу. В случае протекания абсцесса в закрытые полости (плевральную, брюшную) развивается осложнение с гнойным воспалением данных полостей.	Абсцесс имеет вялую клиническую картину и не представляет сложности для диагностики. При глубокой (интратканевой) локализации абсцесса может потребоваться проведение ультразвукового исследования, компьютерной томографии, магнитно-резонансной томографии и других методов - в зависимости от локализации очага в организме.	Лечение абсцесса хирургическое - со вскрытием, опорожнением полости и ее последующим дренированием и промыванием без зашивания. По выражению Н.И.Пирогова «Где гной - там разрез». Обязательно назначаются антибиотики широкого спектра действия.
2. Абсцесс носовой перегородки	Абсцесс носовой перегородки - это гнойный воспалительный процесс, который чаще всего возникает в результате распространения инфекции и гематомы, образовавшейся в результате травмы. Роль причиной заболевания является период гнойной инфекции со стороны нозды, либо гнойной инфекции синусита. Абсцесс сопровождается повышением температуры, нарушением носового дыхания, головной болью.	Для диагностики проводят риноскопию, при которой обнаруживают абсцесс в области перегородки.	Лечение может быть только хирургическим. Проводится вскрытие абсцесса и промывание его полости растворами антисептиков. Назначают системные антибиотики широкого спектра. Местно устанавливают тампоны с антибиотиками и дренажи.
3. Аденоидные разрастания (аденоиды)	Аденоидные разрастания (аденоиды) - это бактериально-воспалительный процесс в лимфоидной ткани носоглотки, который сопровождается увеличением ее размеров, что в свою очередь приводит к затруднению дыхания. Чаще всего болеют дети в возрасте до 7 лет. Предраспространяющиеся фарингиты являются длительные воспалительные процессы слизистой носоглотки при других инфекционных заболеваниях (грипп, ОРВИ, кока и другие). Основными жалобами являются нарушение носового дыхания и выделение из носа. У детей появляется скрип во время сна, возможно повышение слуха. Важным признаком является наличие постоянного гнусавого дыхания. При больших разрастаниях для детей характерны особые виды - с постоянно открытым ртом (аденоидное выражение лица).	Диагностика основана на характерной клинической картине и результатах фарингоскопии и риноскопии.	Лечение может быть консервативным, заключается в назначении противовоспалительных препаратов, промывании носоглотки растворами антисептиков, лазеротерапии, назначении местных глюкокортикоидных мазей. Но единственно эффективным методом лечения является операция удаления аденоидов (аденоидэктомия).
4. Аденоиды простаты	Аденома предстательной железы (простаты) - это доброкачественное новообразование, которое развивается в ней ткани. Развитие образования связано с изменением баланса мужских и женских половых гормонов в мужском организме, которое происходит с возрастом.	Диагноз ставится на основании жалоб пациента, его возраста, данных ректального обследования простаты. Также большим значением обладает анализ мочи, биохимический анализ крови (уровень креатинина и мочевины), проводят определение уровня простатического специфического антигена (PSA). Из инструментальных методов проводят ультразвуковое исследование (УЗИ) простаты на предмет выявления самой аденомы, а также остаточной мочи в мочевого пузыре.	На ранних стадиях аденомы возможно проведение медикаментозного лечения, которое заключается в назначении препаратов, снимающих тонус гладкой мускулатуры (альфа-адреноблокаторы). При доказанном снижении уровня андрогенов назначается андрогенотерапия. В случаях, когда происходит сильное сужение уретры и значительно нарушается нормальное мочеиспускание, прибегают к хирургическому лечению, которое заключается в радикальном удалении простаты. Однако в последнее время наиболее популярным хирургическим методом является малоинвазивная трансуретральная резекция (ТУР) аденомы простаты с помощью эндоскопического и лазерохирургического оборудования.
5. Аденоидиты	Аденоидитальный синдром - это патологическое состояние организма, обусловленное дисфункцией носоглоточных, глоточных с тонзиллярной миндалин и глоточных миндалин (инфекционной инвазивностью). Данная патология чаще встречается у детей. Причины возникновения аденоидитального синдрома следующие: наследственный дефект, характеризующийся недостаточностью фибрилярных систем в носоглоточной полости; врожденный аденоидитальный синдром, который диагностируется сразу после рождения ребенка; увеличение носоглоточных, вследствие опухолевых процессов, гипертрофия носоглоточных. Заболевание носит приобретенный характер. По характеру фибрилярной недостаточности, также различают простую (акрильную) и солитарную формы аденоидитального синдрома. Клиническая картина врожденной формы синдрома проявляется сразу после рождения и характеризуется нарушением строения наружных половых органов у девочек. В некоторых случаях, визуально очень трудно определить пол ребенка. В последующем, такие дети отличаются усиленным ростом, нарастающим мышечной массы, голос становится грубым, отмечается избыточное оволосение (гирсутизм) по мужскому типу, нарушения менструального цикла, либо отсутствие менструаций, отсутствие роста молочных желез и недоразвитие матки и придатков, что приводит к бесплодию. У мальчиков отмечается преждевременное половое развитие, недоразвитие яичек. Отсутствует сперматогенез. Солитарная форма характеризуется сыпчатыми, рыхлыми, дряблыми, дети плохо набирают вес, цвет образования организма, синеватый	Диагностически значимое повышение в крови и в моче уровня гормонов надпочечников. В моче повышается количество 17-оксистероидов. При солитарной форме отмечается повышение содержания калия и снижение натрия.	Медикаментозная терапия, гормональная терапия для подавления гиперактивности аденоидитального гормона, заместительная терапия минералокортикоидами при низком уровне натрий. Санитарно-гигиенические мероприятия: санобработка, витаминные, протодиарейные препараты, биопрепараты. Хирургическое лечение опухоли. Реконструктивная хирургия для формирования правильного типа половых органов.

Рисунок 5 – Фрагмент данных для загрузки в систему

Все использованные исходные данные не удастся привести полностью в связи с их большим объемом, поэтому на рисунке 5 они приведены частично. Данный файл и будет использоваться как исходные данные для системы Эйдос.

### 3.2. Разработка системно-когнитивных моделей и численное решение задач

Ввод исходных данных из таблицы фрагмент которой изображен на рисунке 5 в систему осуществляется с помощью API с параметрами, приведенными на рисунках 6,7.

При загрузке данных были выбраны некоторые ключевые параметры, такие как элементы значений полей-слова символов равным 3, чтобы в результирующую выборку попадали только слова, и проводить лемматизацию, для приведения каждого слова к его лемме, то есть первоначальному состоянию, результаты лемматизации приведены на рисунке 9.

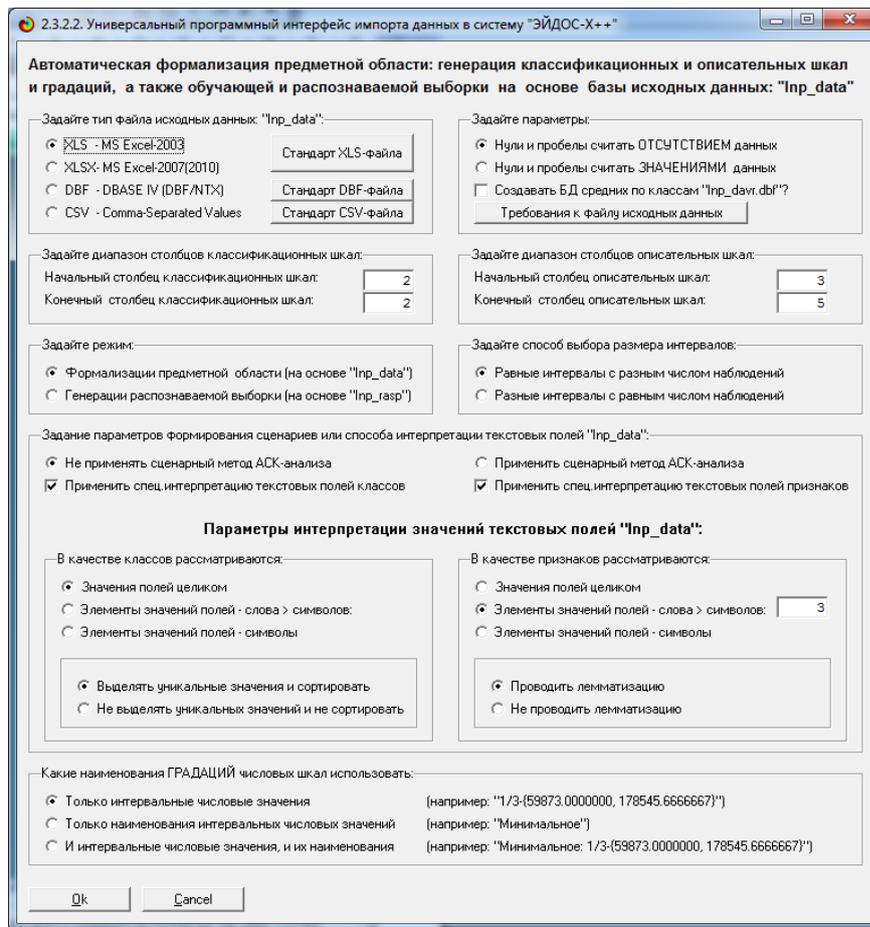


Рисунок 6 – Форма настройки параметров ввода данных

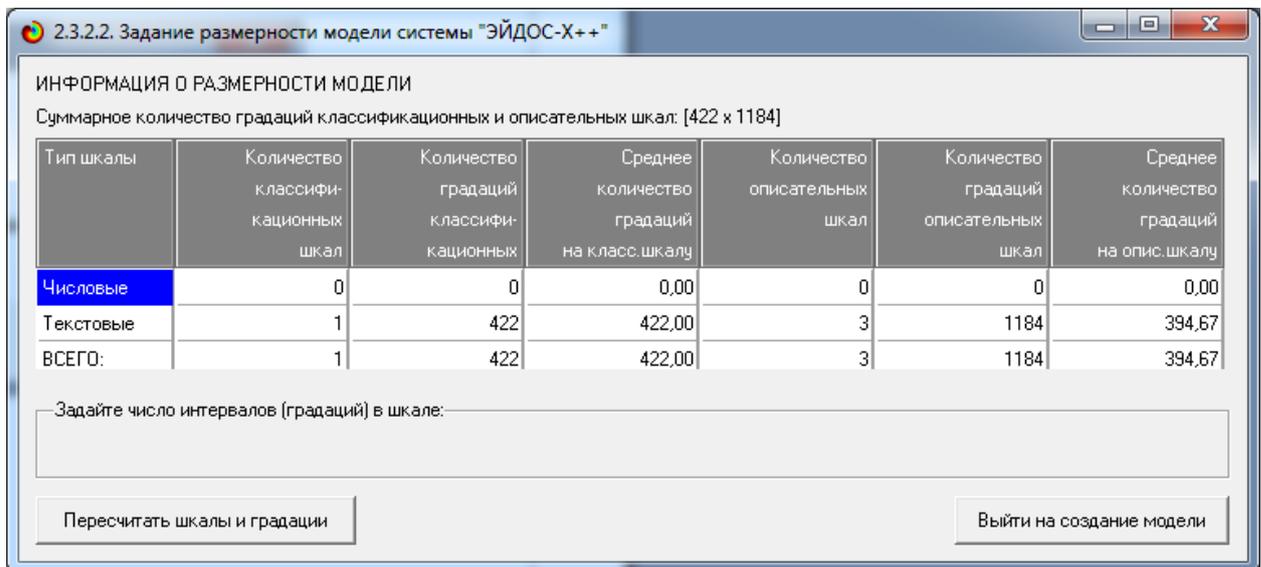


Рисунок 7 – Задание размерности модели системы Эйдос

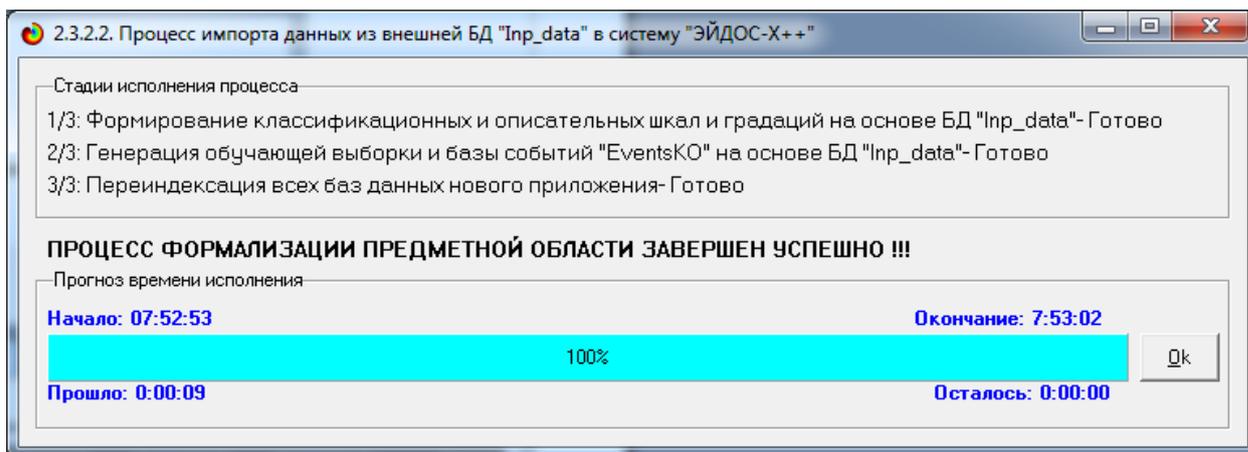


Рисунок 8 – Процесс импорта данных

На рисунке 8 изображен процесс импорта данных из файла Inp\_data

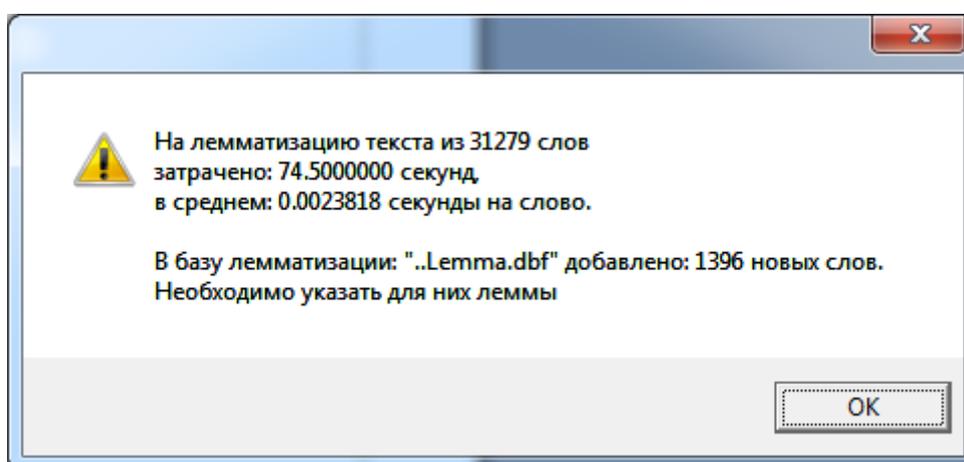


Рисунок 9 – Лемматизация исходных данных

При загрузке данных были выбраны некоторые ключевые параметры, такие как элементы значений полей-слова символов равным 3, чтобы в результирующую выборку попадали только слова, и проводить лемматизацию, для приведения каждого слова к его лемме, то есть первоначальному состоянию, результаты лемматизации приведены на рисунке 9.

Кроме самого ввода исходных данных API осуществляет также формализацию предметной области, которая включает разработку классификационных и описательных шкал и градаций (таблица 6, таблица 7), также вид в программном интерфейсе (рисунок 10), а затем и кодирование исходных

данных с их помощью, в результате чего они преобразуются в обучающую выборку (рисунок 11).

Таблица 6 – Классификационные и градации – вербальные нозологические образы (фрагмент)

Код	Наименование
1	ДИАГНОЗ-Абсцесс мягких тканей
2	ДИАГНОЗ-Абсцесс носовой перегородки
3	ДИАГНОЗ-Аденоидные разрастания (аденоиды)
4	ДИАГНОЗ-Аденома простаты
5	ДИАГНОЗ-Адреногенитальные расстройства
6	ДИАГНОЗ-Акромегалия
7	ДИАГНОЗ-Алкогольная полинейропатия
8	ДИАГНОЗ-Алкогольный гепатит
9	ДИАГНОЗ-Аллергия
10	ДИАГНОЗ-Алопеция
11	ДИАГНОЗ-Альдостерома
12	ДИАГНОЗ-Амиотрофический боковой склероз
13	ДИАГНОЗ-Анальная трещина
14	ДИАГНОЗ-Ангина
15	ДИАГНОЗ-Ангина Симановского-Венсана (язвенно-некротическая)
16	ДИАГНОЗ-Аневризма брюшной аорты
17	ДИАГНОЗ-Анемия
18	ДИАГНОЗ-Анкилоз коленного сустава
19	ДИАГНОЗ-Анкилоз локтевого сустава
20	ДИАГНОЗ-Аномалии развития наружного и среднего уха
21	ДИАГНОЗ-Аппендицит
22	ДИАГНОЗ-Аскаридоз
23	ДИАГНОЗ-Астигматизм
24	ДИАГНОЗ-Атопический дерматит (нейродермит)
25	ДИАГНОЗ-Аутоиммунный тиреоидит
26	ДИАГНОЗ-Ахиллобурсит
27	ДИАГНОЗ-Бедренная грыжа
28	ДИАГНОЗ-Билиарный цирроз печени
29	ДИАГНОЗ-Болезнь Альцгеймера
30	ДИАГНОЗ-Болезнь Бехтерева
31	ДИАГНОЗ-Болезнь Гиршпрунга

Таблица 6 – Классификационные и градации – вербальные нозологические образы (продолжение)

32	ДИАГНОЗ-Болезнь Кенига
33	ДИАГНОЗ-Болезнь Крона
34	ДИАГНОЗ-Болезнь Меньера
35	ДИАГНОЗ-Болезнь Пика
36	ДИАГНОЗ-Ботулизм
37	ДИАГНОЗ-Бронхиальная астма
38	ДИАГНОЗ-Бронхоэктатическая болезнь
39	ДИАГНОЗ-Брюшной тиф
40	ДИАГНОЗ-Бурсит
41	ДИАГНОЗ-Варикозное расширение вен
42	ДИАГНОЗ-Варикоцеле
43	ДИАГНОЗ-Ветряная оспа
44	ДИАГНОЗ-Внутрисуставные переломы плечевой кости
45	ДИАГНОЗ-Внутрисуставные тела коленного сустава
46	ДИАГНОЗ-Водянка оболочек яичка
47	ДИАГНОЗ-Врожденные деформации наружного носа и носовой перегородки
48	ДИАГНОЗ-Врожденные кисты и свищи носа
49	ДИАГНОЗ-Врожденные околоушные свищи и кисты
50	ДИАГНОЗ-Врожденные пороки развития наружного и среднего уха
51	ДИАГНОЗ-Врожденные пороки сердца
52	ДИАГНОЗ-Врожденный вывих бедра, остаточный подвывих бедра

Таблица 7 – Описательные шкалы и градации – слова (фрагмент, всего в модели 6461 слово)

Код	Наименование
1	СИМПТОМЫ-chlamydia
2	СИМПТОМЫ-herpesviridae
3	СИМПТОМЫ-paramyxovirus
4	СИМПТОМЫ-parapertusis
5	СИМПТОМЫ-plasmodium
6	СИМПТОМЫ-simplex
7	СИМПТОМЫ-trahomatis
8	СИМПТОМЫ-абсолютный
9	СИМПТОМЫ-абсцесс
10	СИМПТОМЫ-авария

Таблица 7 – Описательные шкалы и градации – слова (продолжение)

Код	Наименование
11	СИМПТОМЫ-автоматический
12	СИМПТОМЫ-автомобильный
13	СИМПТОМЫ-автор
14	СИМПТОМЫ-аддисонический
15	СИМПТОМЫ-аденоид
16	СИМПТОМЫ-аденоидный
17	СИМПТОМЫ-аденокарцинома
18	СИМПТОМЫ-аденокарциномой
19	СИМПТОМЫ-аденома
20	СИМПТОМЫ-адреналин
21	СИМПТОМЫ-адреногенитальный
22	СИМПТОМЫ-адренокортикотропного
23	СИМПТОМЫ-азотистый
24	СИМПТОМЫ-акромегалия
25	СИМПТОМЫ-актг
26	СИМПТОМЫ-активность
27	СИМПТОМЫ-активный
28	СИМПТОМЫ-акустический
29	СИМПТОМЫ-алкоголь
30	СИМПТОМЫ-алкогольный
31	СИМПТОМЫ-аллерген
32	СИМПТОМЫ-аллергический
33	СИМПТОМЫ-аллергия
34	СИМПТОМЫ-алопеции
35	СИМПТОМЫ-алопеция
36	СИМПТОМЫ-альвеола
37	СИМПТОМЫ-альдостерома
38	СИМПТОМЫ-альдостеромы
39	СИМПТОМЫ-альцгеймера
40	СИМПТОМЫ-аминокислота
41	СИМПТОМЫ-амиотрофический
42	СИМПТОМЫ-амниотический
43	СИМПТОМЫ-анастомоза
44	СИМПТОМЫ-анальный
45	СИМПТОМЫ-анатомический
46	СИМПТОМЫ-ангина
47	СИМПТОМЫ-ангиофиброма

Таблица 7 – Описательные шкалы и градации – слова (продолжение)

Код	Наименование
48	СИМПТОМЫ-андроген
49	СИМПТОМЫ-аневризм
50	СИМПТОМЫ-аневризма
51	СИМПТОМЫ-анемия
52	СИМПТОМЫ-анкилоз

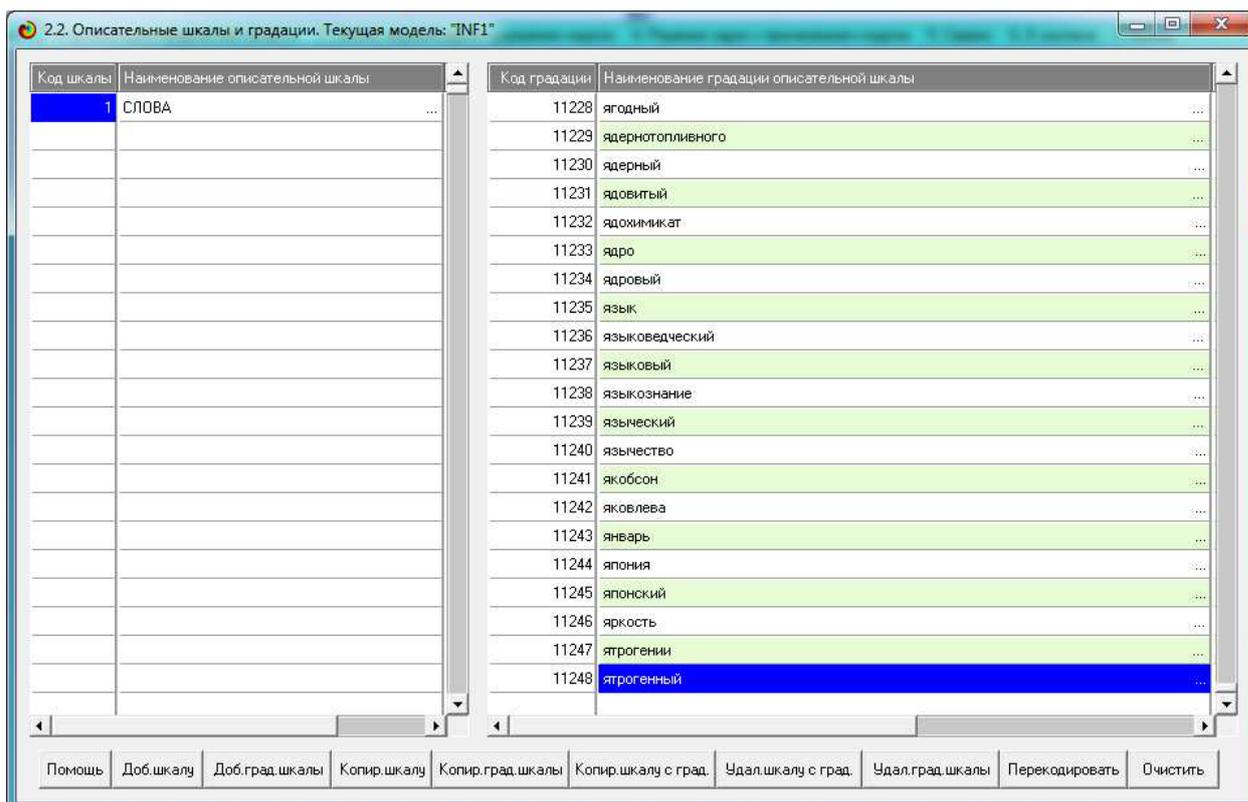


Рисунок 10 - Описательные шкалы и градации вид в программном интерфейсе – слова (фрагмент, всего в модели 6461 слово)

Обучающая выборка представляет собой базу исходных данных, закодированную с помощью классификационных и описательных шкал и градаций (рисунок 11).

На рисунке 11 приведен только фрагмент обучающей выборки, т.к. в нее входит 6461 слов. В верхнем окне рисунка обучающей выборки приведены наименования объектов обучающей выборки, в левой части – коды классов, связанных с этими объектами, в правой части – коды признаков, которые имеются у этих объектов.

The screenshot shows a software window titled "2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: 'INF5'". It contains two data tables and a control panel at the bottom.

**Table 1: Object Information**

Код объекта	Наименование объекта	Дата	Время
1	Абсцесс мягких тканей	...	
2	Абсцесс носовой перегородки	...	
3	Аденоидные разрастания (аденоиды)	...	
4	Аденома простаты	...	
5	Аденогенитальные расстройства	...	
6	Акромегалия	...	
7	Алкогольная полинейропатия	...	
8	Алкогольный гепатит	...	
9	Аллергия	...	
10	Аллергия	...	

**Table 2: Feature Data**

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4	Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	1	0	0	0	1	9	407	1220	1518	409	252	914
					1	1258	2083	1559	994	153	1568	913
					1	903	225	1217	2377	2021	2007	762
					1	1345	2396	2834	4160	2945	3559	3544
					1	3785	2678	2554	3047	2396	3122	3518
					1	3580	3988	2875	2969	3929	4171	6328
					1	4446	5377	5598	5627	4670	5744	6273
					1	4474	5513	4538	5815	5344	5184	4233

**Control Panel:** Помощь | Скопировать обуч.выб.в расп. | Добавить объект | Добавить классы | Добавить признаки | Удалить объект | Удалить классы | Удалить признаки | Очистить БД

Рисунок 11 – Обучающая выборка (фрагмент)

Таким образом созданы представления нормализованных исходных данных.

Обучающая выборка представляет все необходимые и достаточные условия для проведения следующего этапа АСК-анализа. Следующий этап – синтез и верификация моделей.

На рисунке 12 представлен предлагаемый интерфейс режима синтеза и верификации моделей:

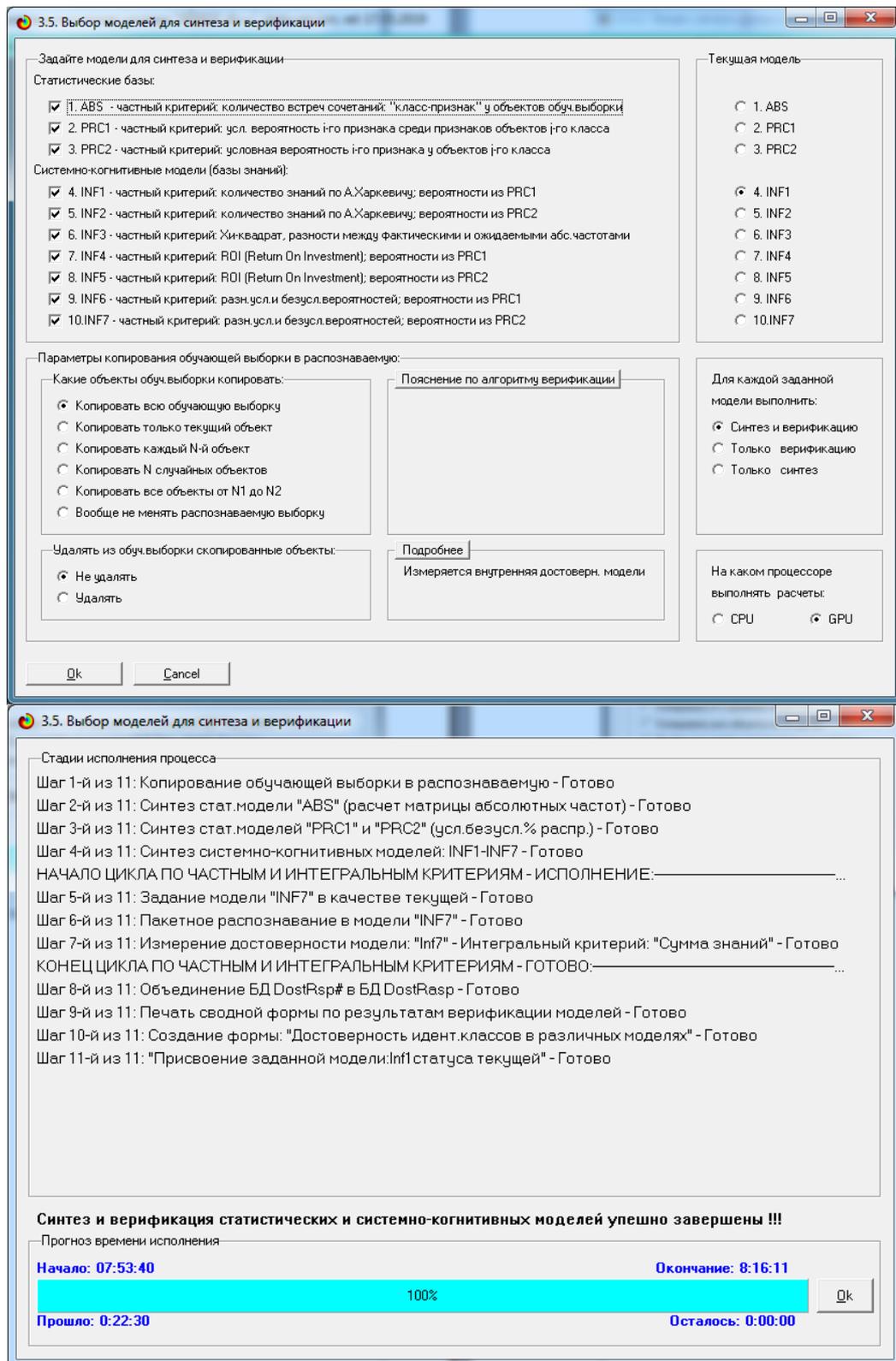


Рисунок 12 – Предлагаемые экранные формы режима синтеза и верификации моделей и стадии процесса исполнения

На рисунке 12 видно, что весь процесс синтеза и верификации моделей занял 22 секунды. Отметим, что при синтезе и верификации был использован процессор графической видеокарты (GPU). На основном процессоре (CPU) выполнение этих операций занимает значительно больше времени. На некоторых задачах время увеличивается в несколько десятков раз). Программа позволяет максимально эффективно использовать вычислительные ресурсы компьютера.

Отметим также, что база лемматизации А.А.Зализняка включает 2097005 слов. В ходе выполнения данной работы она была дополнена еще на 1396, представляющих собой термины, встретившиеся в определениях медицинских болезней. Сами модели имеют вид, приведенный на рисунке 13:

Код признака	Наименование описательной шкалы и градации	1. ДИАГНОЗ АБСЦЕСС МЯГКИХ ТКАНЕЙ	2. ДИАГНОЗ АБСЦЕСС НОСОВОЙ ПЕРЕГОРОДКИ	3. ДИАГНОЗ АДЕНОИДНЫЕ РАЗРАСТАНИЯ (АДЕНОИДЫ)	4. ДИАГНОЗ АДЕНОМА ПРОСТАТЫ	5. ДИАГНОЗ АДРЕНОГЕН... РАССТРОЙСТ...	6. ДИАГНОЗ АКРОМЕГАЛИЯ	7. ДИАГНОЗ АЛКОГОЛЬНАЯ ПОЛИНЕЙРО...	8. ДИАГНОЗ АЛКОГОЛЬН... ГЕПАТИТ	9. ДИАГНОЗ АЛЛЕРГИЯ	10. ДИАГНОЗ АЛОПЕЦИЯ	11. ДИАГНОЗ АЛЬДОСТЕР...
2079	СИМПТОМЫ типичный	-0.009	-0.008	-0.007	-0.011	-0.009	-0.010	-0.009	-0.009	-0.010	-0.007	-0.010
2080	СИМПТОМЫ тиреоидит	-0.009	-0.008	-0.007	-0.011	-0.009	-0.010	-0.009	-0.009	-0.010	-0.007	-0.010
2081	СИМПТОМЫ тиреоидные	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
2082	СИМПТОМЫ тиреотокс	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
2083	СИМПТОМЫ ткань	0.947	-0.048	0.957	0.939	-0.052	0.943	-0.053	-0.053	-0.056	-0.041	-0.055
2084	СИМПТОМЫ тканьё	-0.033	-0.029	-0.026	-0.037	-0.031	-0.035	-0.033	-0.033	-0.034	-0.025	-0.034
2085	СИМПТОМЫ тога	-0.007	-0.006	-0.006	-0.008	-0.007	-0.007	-0.007	-0.007	-0.007	-0.005	-0.007
2086	СИМПТОМЫ токсикогенный	-0.005	-0.004	-0.004	-0.005	-0.004	-0.005	-0.005	-0.005	-0.005	-0.004	-0.005
2087	СИМПТОМЫ токсикоз	-0.014	-0.012	-0.011	-0.016	-0.013	-0.015	-0.014	-0.014	-0.015	-0.011	-0.014
2088	СИМПТОМЫ токсин	-0.005	-0.004	-0.004	-0.005	-0.004	-0.005	-0.005	-0.005	-0.005	-0.004	-0.005
2089	СИМПТОМЫ токсический	-0.014	-0.012	-0.011	-0.016	-0.013	-0.015	-0.014	0.986	-0.015	-0.011	-0.014
2090	СИМПТОМЫ токсоплазм	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
2091	СИМПТОМЫ толстый	-0.009	-0.008	-0.007	-0.011	-0.009	-0.010	-0.009	-0.009	-0.010	-0.007	-0.010
2092	СИМПТОМЫ толщенный	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
2093	СИМПТОМЫ только	-0.023	-0.021	-0.019	-0.027	-0.022	-0.025	-0.023	-0.023	-0.025	-0.018	-0.024
2094	СИМПТОМЫ тонзиллит	-0.007	-0.006	-0.006	-0.008	-0.007	-0.007	-0.007	-0.007	-0.007	-0.005	-0.007
2095	СИМПТОМЫ тонкий	-0.009	-0.008	-0.007	-0.011	-0.009	-0.010	-0.009	-0.009	-0.010	-0.007	-0.010
2096	СИМПТОМЫ тонкостенный	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
2097	СИМПТОМЫ тонкий	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002

Рисунок 13 – Системно-когнитивная модель INF3 (фрагмент)

Проведена оценка надежности моделей в системе «Эйдос» путем решения проблемы классификации исходных изображений в соответствии с изображением классов и подсчета количества истинно положительных и от-

рицательных, а также ложных положительных и отрицательных решений по F-мере Ван-Ризбергена и L1-L2-мер проф. Е.В. Луценко.

Классическая количественная мера достоверности моделей: F-мера Ван Ризберген основан на подсчете общего количества истинных и ошибочно идентифицированные и неопознанные объекты образец в мульти классовых системах классификации, таких как система «Эйдос», тот же обучающий или узнаваемый объект образцы могут одновременно применяться ко многим классам. В соответствии с

При синтезе модели ее описание используется для формирования обобщенные образы многих классов, к которым он принадлежит. Когда используется Использование модели для классификации определяется степенью сходства различия объекта со всеми классами, и истинно-положительные Это может быть объект, принадлежащий нескольким классам одновременно себя. В результате этой классификации оказывается, что объект не просто правильно или ошибочно относится или не относится к разным сам, как в классической F-мере, но правильно или неправильно ссылается или не распространяется на них в разной степени.

Однако классическая F-мера не учитывает тот факт, что объект может на самом деле, он одновременно принадлежит ко многим классам (мульти класс тот факт, что в результате классификации, разнообразие личностная степень сходства, различия объекта с классами (нечеткость). На многочисленные численные примеры автора обнаружили, что при использовании мутные положительные и истинно отрицательные решения, модуль подобия различия между объектом и классами значительно выше, чем при ложном положительные и ложноотрицательные решения.

Поэтому была предложена мера надежности модели L1, принимая во внимание не только тот факт, является истинным или ложным положительным или отрицательное решение, но и степень доверия классификатора к этим решениям. Однако при классификации больших данных большое коли-

чество ложноположительных решений с низким сходством, которые в целом могут внести большой вклад в обоснованность модели.

Чтобы преодолеть эту проблему, предложена L2-мера, в которой вместо сумм уровней сходства с использованием средних уровней сходства согласно различным вариантам классификации.

Таким образом, в системе «Эйдос» меры принимаются меры достоверности модели, называемые L1-мера и L2-мера, которые смягчают и преодолевают недостатки F-меры. В работе эти меры описаны математически и их приложение демонстрируется на численном примере. В интеллектуальной системе «Эйдос», представляющей собой программный инструмент автоматизированный системно-когнитивный анализ (ASC-анализ), все эти меры доверия модели расположены: F, L1 и L2

В режиме 4.1.3.6 кратко и в режиме 4.1.3.7 показано более подробно надежность каждой конкретной модели в соответствии с этими мерами. В этом случае мера L2 имеет наивысшую точность. В соответствии с критерием  $L2=0.997$  наиболее достоверной является модель INF5 (рисунок 14):

Наименование модели и частного критерия	Интегральный критерий	Средняя популяционная оценка ложно-отриц. решений (SF)	S-T точность модели	S-Помога модели	L1 мера прогн. Э.В. Ляденко	Средний популяционный сход. истинно-положит. решений	Средний популяционный сход. ложно-положит. решений	Средний популяционный сход. ложно-отриц. решений	Средний популяционный сход. ложно-положит. решений	А-Т точность модели AР/Reson = ATR/AATP	А-Помога модели AР/Reson = ATR/AATP	L2-мера прогн. Э.В. Ляденко
1. ABS - частный критерий: количество встреч сонетный "клас...	Корреляция абс. частот с обр...	0,049	1,000	0,094	1,000	0,008	0,051	0,952	1,000	0,975	0,975	0,975
1. ABS - частный критерий: количество встреч сонетный "клас...	Средня абс. частот по призна...	0,044	1,000	0,094	0,707	0,008	0,041	0,946	1,000	0,972	0,972	0,972
2. PRC1 - частный критерий: усл. вероятность его признава сред...	Корреляция усл. от частот с о...	0,049	1,000	0,094	1,000	0,008	0,051	0,952	1,000	0,975	0,975	0,975
2. PRC1 - частный критерий: усл. вероятность его признава сред...	Средня усл. от частот по при...	0,043	1,000	0,093	0,676	0,008	0,039	0,945	1,000	0,972	0,972	0,972
3. PRC2 - частный критерий: условная вероятность его признава...	Корреляция усл. от частот с о...	0,049	1,000	0,094	1,000	0,008	0,051	0,952	1,000	0,975	0,975	0,975
3. PRC2 - частный критерий: условная вероятность его признава...	Средня усл. от частот по при...	0,044	1,000	0,094	0,707	0,008	0,041	0,946	1,000	0,972	0,972	0,972
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0,106	1,000	0,192	0,952	0,005	0,025	0,975	1,000	0,987	0,987	0,987
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Средня знаний	0,083	1,000	0,153	0,757	0,005	0,022	0,972	1,000	0,986	0,986	0,986
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0,106	1,000	0,192	0,952	0,005	0,024	0,975	1,000	0,987	0,987	0,987
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Средня знаний	0,083	1,000	0,154	0,729	0,005	0,021	0,972	1,000	0,986	0,986	0,986
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	0,140	1,000	0,248	0,989	0,007	0,036	0,965	1,000	0,982	0,982	0,982
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Средня знаний	0,143	1,000	0,250	0,699	0,019	0,025	0,965	1,000	0,982	0,982	0,982
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0,437	1,000	0,608	0,783	0,005	0,012	0,985	1,000	0,993	0,993	0,993
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Средня знаний	0,298	1,000	0,459	0,500	0,003	0,003	0,994	1,000	0,997	0,997	0,997
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0,437	1,000	0,608	0,783	0,005	0,012	0,985	1,000	0,993	0,993	0,993
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Средня знаний	0,300	1,000	0,461	0,494	0,003	0,003	0,994	1,000	0,997	0,997	0,997
9. INF6 - частный критерий: разности между вероятностей; вер...	Семантический резонанс зна...	0,062	1,000	0,116	0,995	0,007	0,041	0,960	1,000	0,980	0,980	0,980
9. INF6 - частный критерий: разности между вероятностей; вер...	Средня знаний	0,053	1,000	0,100	0,676	0,006	0,032	0,955	1,000	0,977	0,977	0,977
10. INF7 - частный критерий: разности между вероятностей; ве...	Семантический резонанс зна...	0,062	1,000	0,116	0,995	0,006	0,041	0,960	1,000	0,980	0,980	0,980
10. INF7 - частный критерий: разности между вероятностей; ве...	Средня знаний	0,053	1,000	0,101	0,695	0,006	0,033	0,955	1,000	0,977	0,977	0,977

Рисунок 14. Предлагаемые экранные формы режима верификации моделей

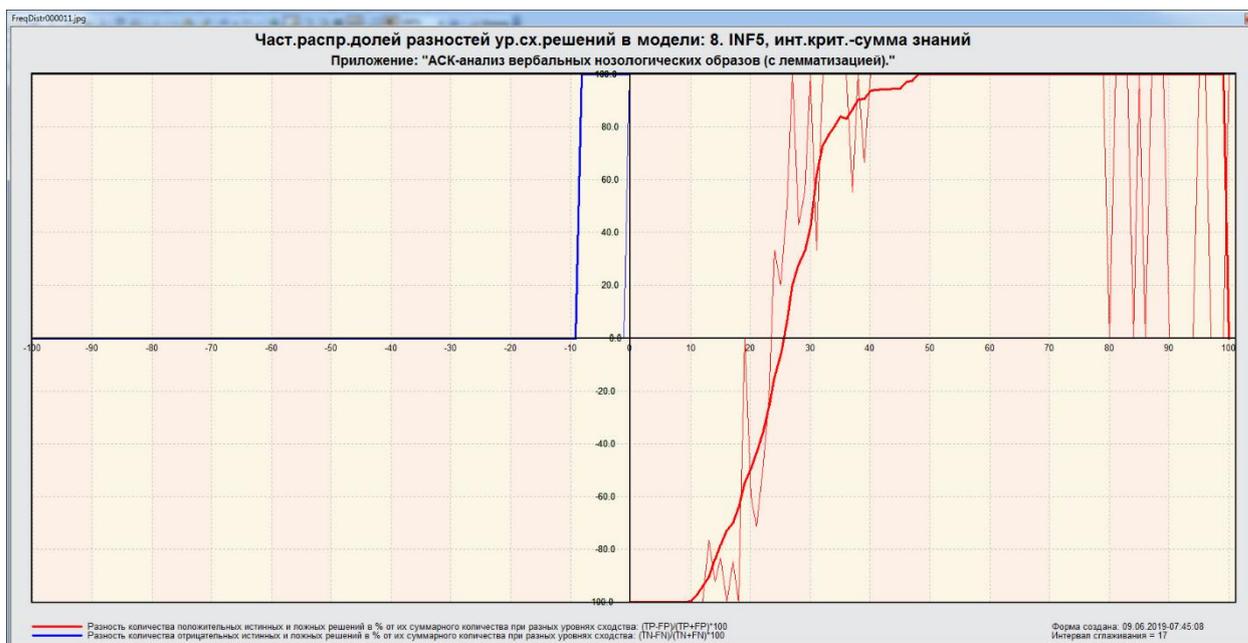


Рисунок 14. Предлагаемые экранные формы режима верификации моделей  
(продолжение)

Из рисунка 14 мы видим, что отрицательные решения (т.е. решения о непринадлежности объекта к классу) в данной модели всегда истинные, а положительные решения (т.е. о принадлежности объекта к классу) есть и истинные, и ложные, причем все ложные решения наблюдаются при невысоких уровнях сходства (примерно до 5%), а более высоких уровнях сходства наблюдаются только истинно-положительные решения. Это позволяет надежно и с высокой достоверностью решать задачи системной классификации и идентификации текстов.

Это очень хорошие и разумные результаты. Обратимся к режиму 3.7.5. Данный режим показывает Парето зависимость суммарной дифференцирующей мощности модели от числа градаций описательных шкал, рассортированных в порядке убывания их селективной силы или значимости (рисунок 15).

В качестве значимости значения фактора для решения различных задач в системно-когнитивной модели (СК-модель) в АСК-анализе и системе «Эй-

дос» принята вариабельность значений фактора по классам в матрице модели.

В качестве значимости фактора для решения различных задач в системно-когнитивной модели (СК-модель) в АСК-анализе и системе «Эйдос» принята средняя значимость значений данного фактора по классам в матрице модели.

Существует много количественных мер вариабельности (изменчивости), но в данном случае принято использовать среднеквадратичное отклонение.

Степень детерминированности вербальных нозологических образов (классов) в наиболее достоверной модели изображена на рисунке 15.

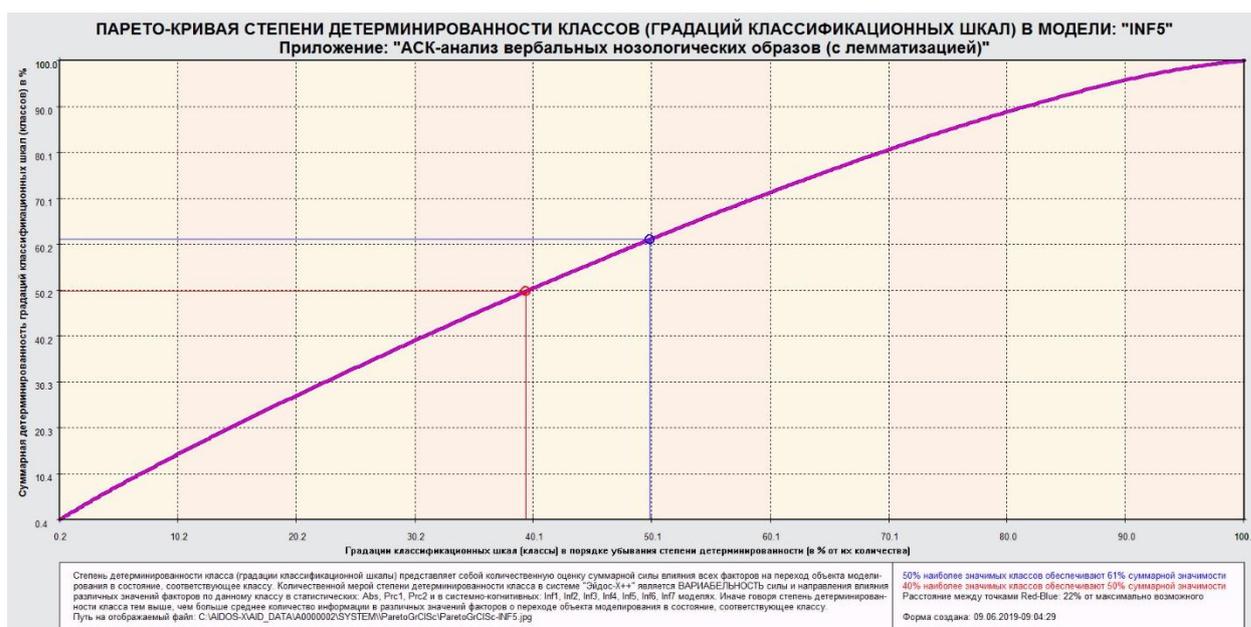


Рисунок 15 – Степень детерминированности вербальных нозологических образов (классов) в модели INF5.

Вариабельность значения фактора по классам выбрана в качестве значимости этого значения фактора потому, что чем выше эта вариабельность, тем лучше позволяет это значение фактора разделить (различить) классы. Если вариабельность значения фактора равна нулю, то данное значение фактора

вообще является бесполезным для решения задачи идентификации и других задач.

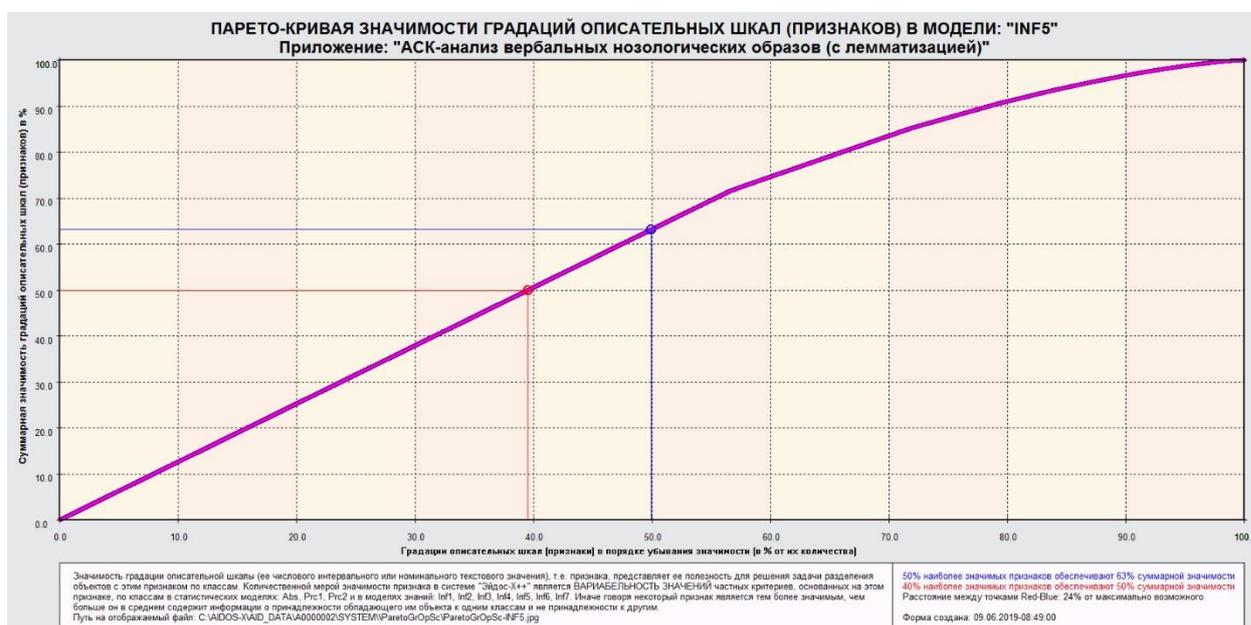


Рисунок 16 – Парето-зависимость суммарной дифференцирующей мощности модели от числа градаций описательных шкал, рассортированных в порядке убывания их селективной силы в СК-модели INF5.

На основе рисунка 16 и соответствующих таблиц можно обоснованно сделать выводы о том, что, например, в наиболее достоверной модели INF5:

1. 50% наиболее значимых градаций описательных шкал обеспечивают 91% суммарной селективной мощности модели (точка BLUE).
2. 50% суммарной селективной мощности модели обеспечивается 5% наиболее значимых градаций описательных шкал (точка RED).
3. Число градаций описательных шкал может быть существенно сокращено без особой потери качества модели путем удаления из модели мало-значимых градаций. При этом размерность модели существенно сократится и ее быстроедействие соответственно возрастет.

Для решения задачи прогнозирования возможных диагнозов с конкретным пациентом необходимо предварительно придать статус текущей модели INF5 (рисунок 17), а уже затем проводить собственно прогнозирование (рисунок 18):

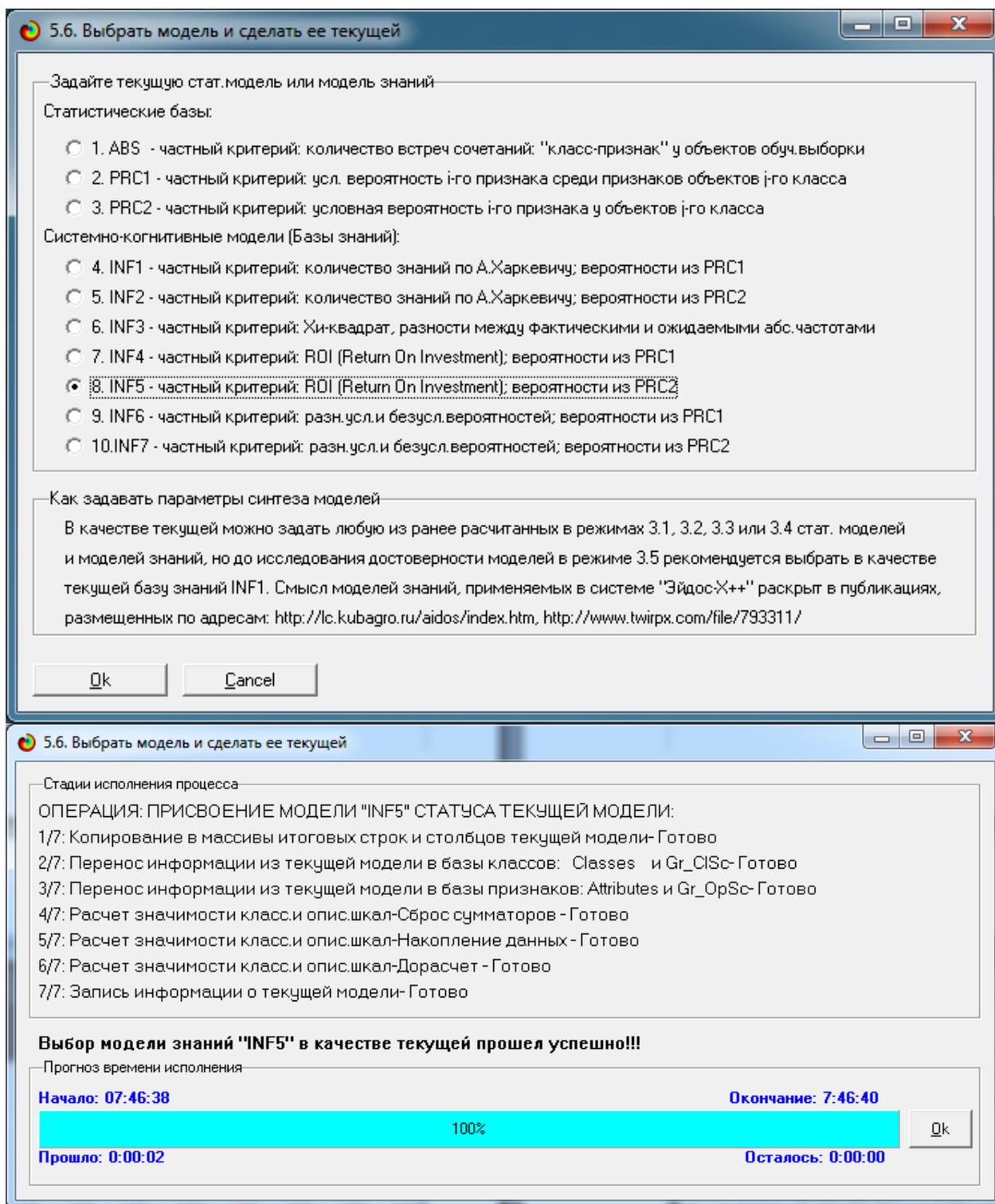


Рисунок 17 – Придание модели INF5 статуса текущей

Экранные формы режима придания наиболее достоверной модели статуса текущей.

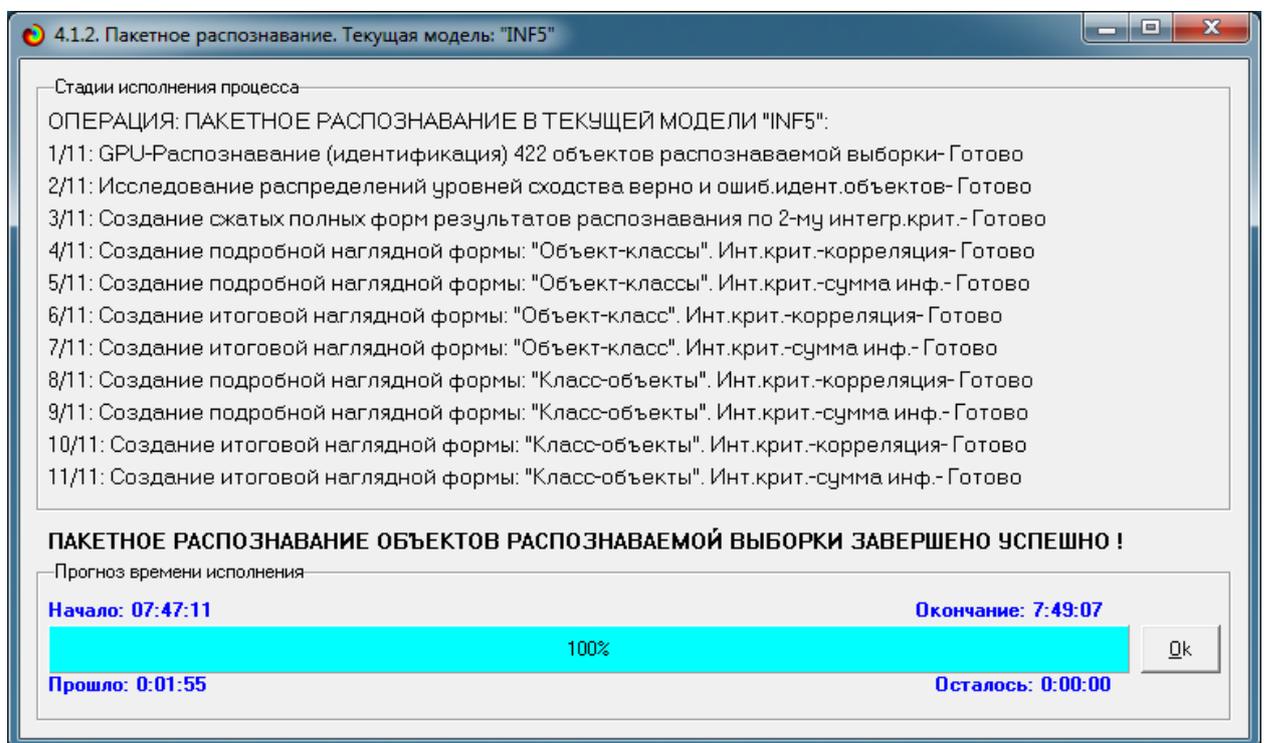
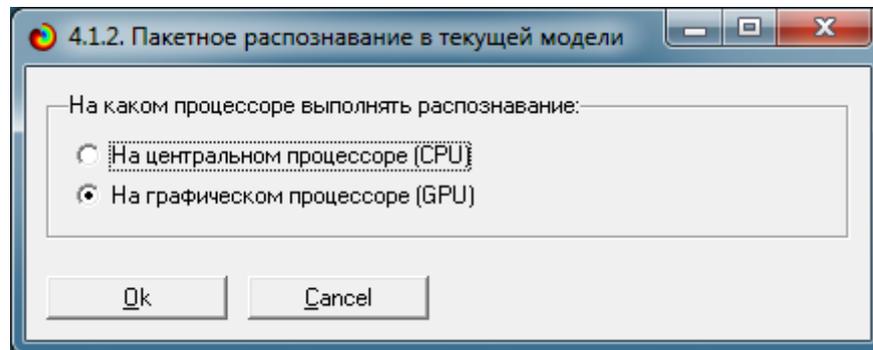


Рисунок 18 – Экранные формы режима решения задачи идентификации в наиболее достоверной модели

Ввод исходных для прогнозирования осуществляется с помощью того же API, что и исходных данных для формирования моделей (рисунок 6), только с опцией: «Генерация распознаваемой выборки».

Видно, что идентификация 82 текстов в ранее созданных моделях заняла 1 минуту 55 секунд. Отметим, что большую часть этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этой идентификации. Эти формы отражают результаты идентификации в различных разрезах и обобщениях.

Результат прогнозирования получаем в виде экранной формы, в которой все возможные результаты ранжированы в порядке убывания релевантности (рисунок 19):

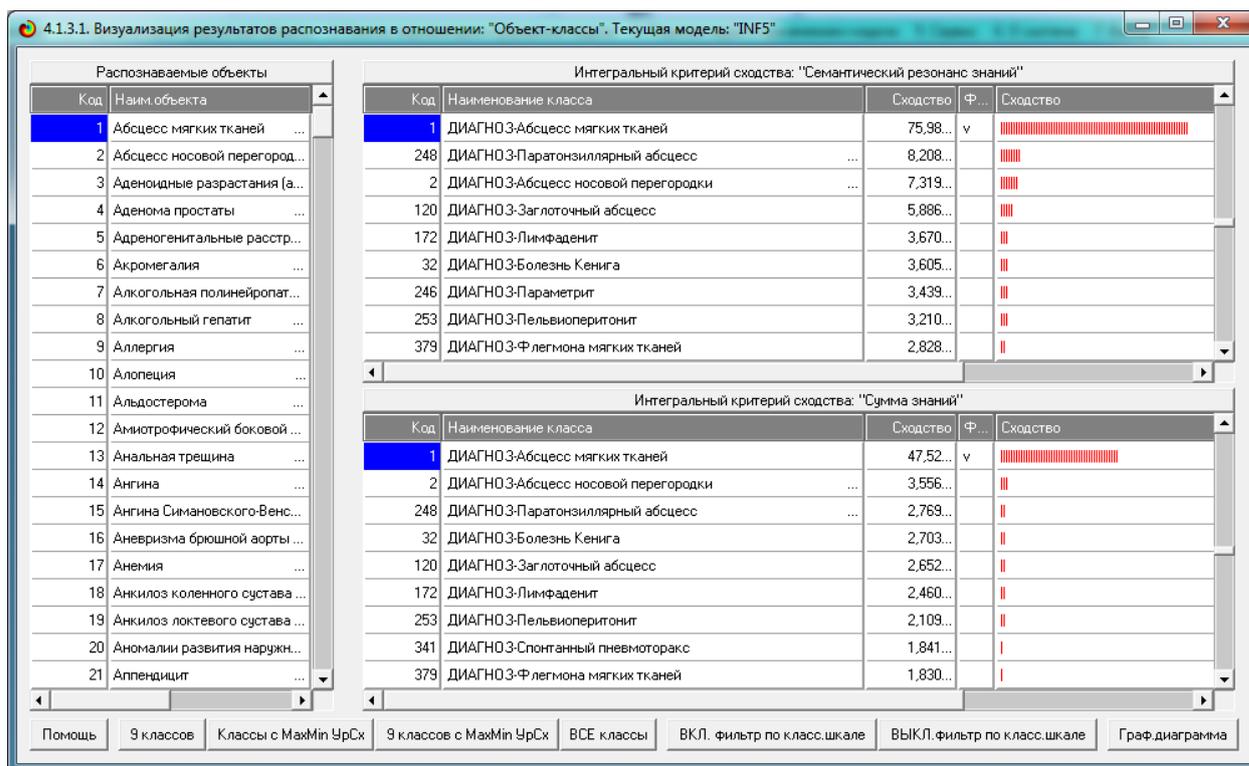


Рисунок 19. Экранная форма прогнозирования результатов диагноза

Видно, что ожидается минимальный результат, чуть менее вероятен вариант немного лучше, и т.д. Синим отмечены варианты, которые согласно прогнозу исключаются.

На подобные прогнозы врач может опираться при установлении диагноза пациента.

Приведем еще две результирующие формы рисунок 20.

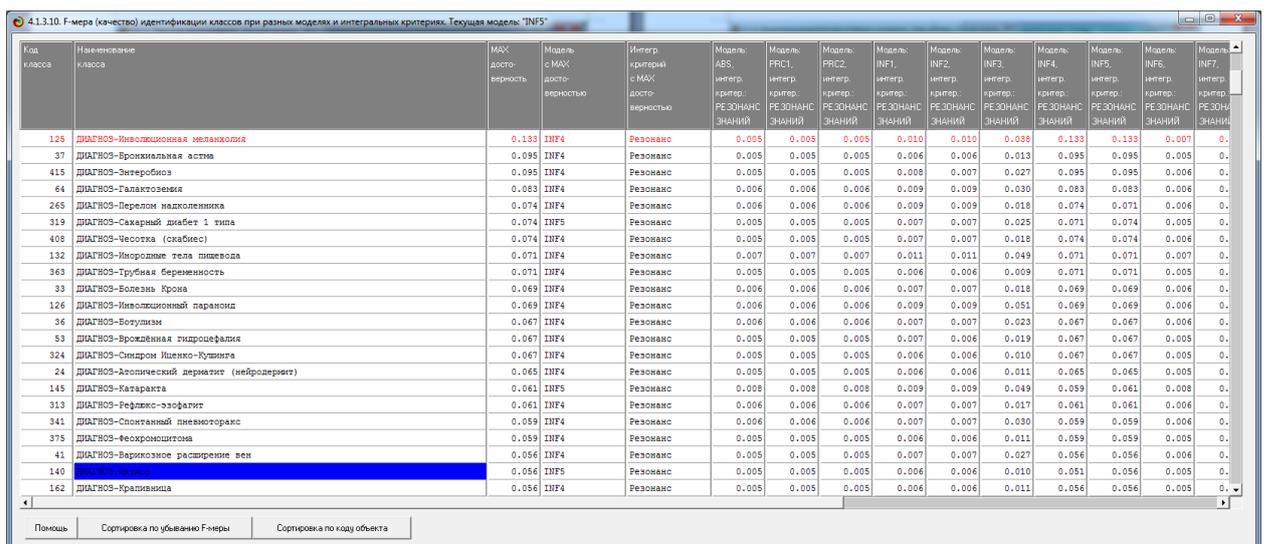
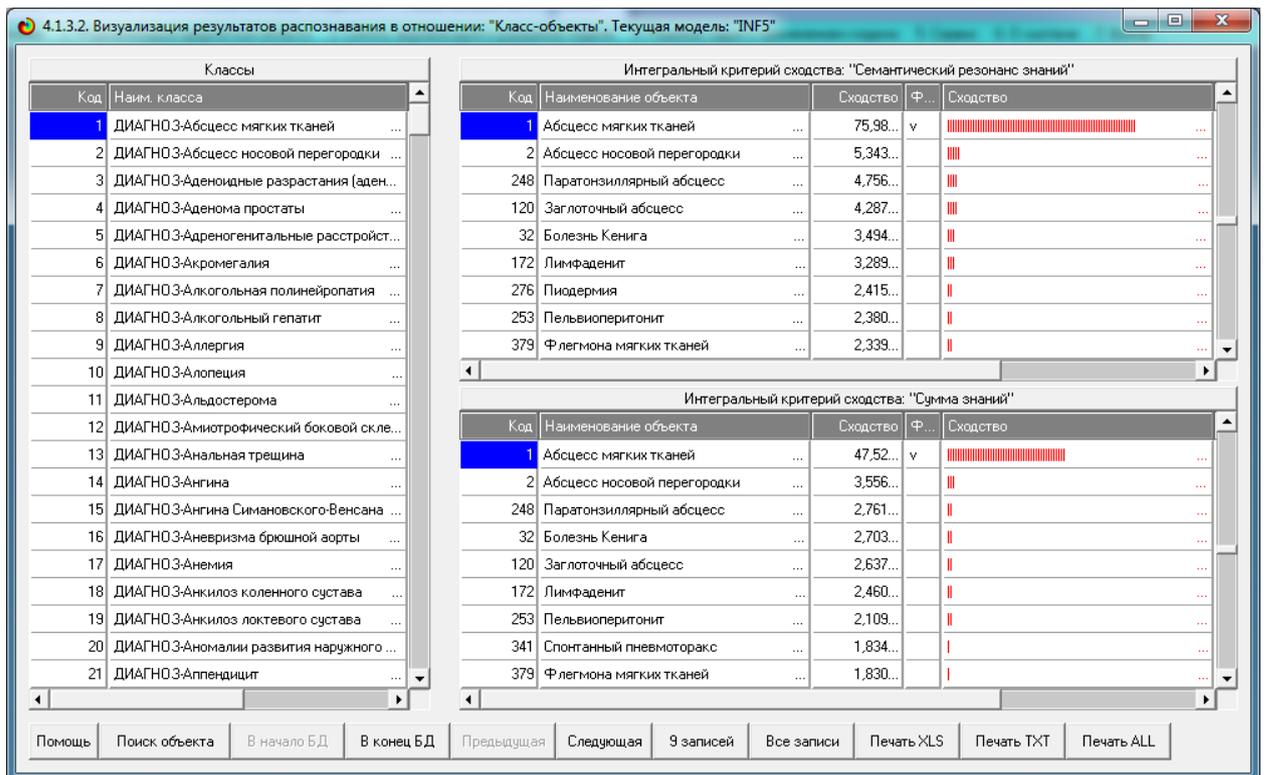


Рисунок 21. Выходные формы по результатам верификации моделей

Из рисунка 21 видно, что результаты атрибуции имеют очень высокую достоверность и полностью соответствуют экспертным ожиданиям.

Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» без фильтра изображены на рисунке 22.

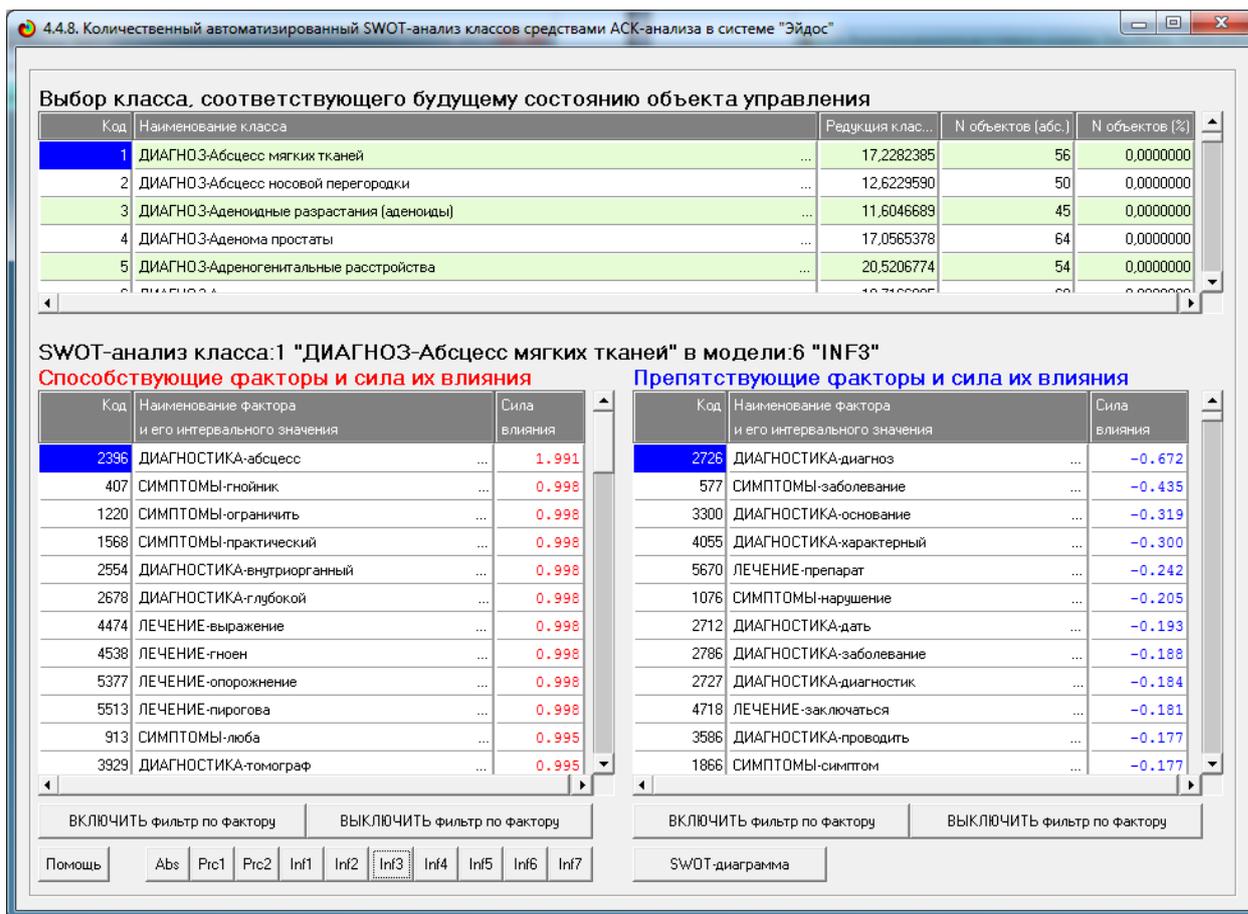


Рисунок 22 - Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (без фильтра)

При принятии решений определяется сила и направления влияния факторов на принадлежность состояний объекта моделирования к тем или иным классам. По сути это решение задачи SWOT-анализа. В системе «Эйдос» в режиме 4.4.8 поддерживается решение этой задачи. При этом выявляется семантическое ядро и семантическое анти ядро заданного класса. На рисунке 23 приведено семантическое ядро и анти ядро нозологического образа диагноза абсцесс мягких тканей.

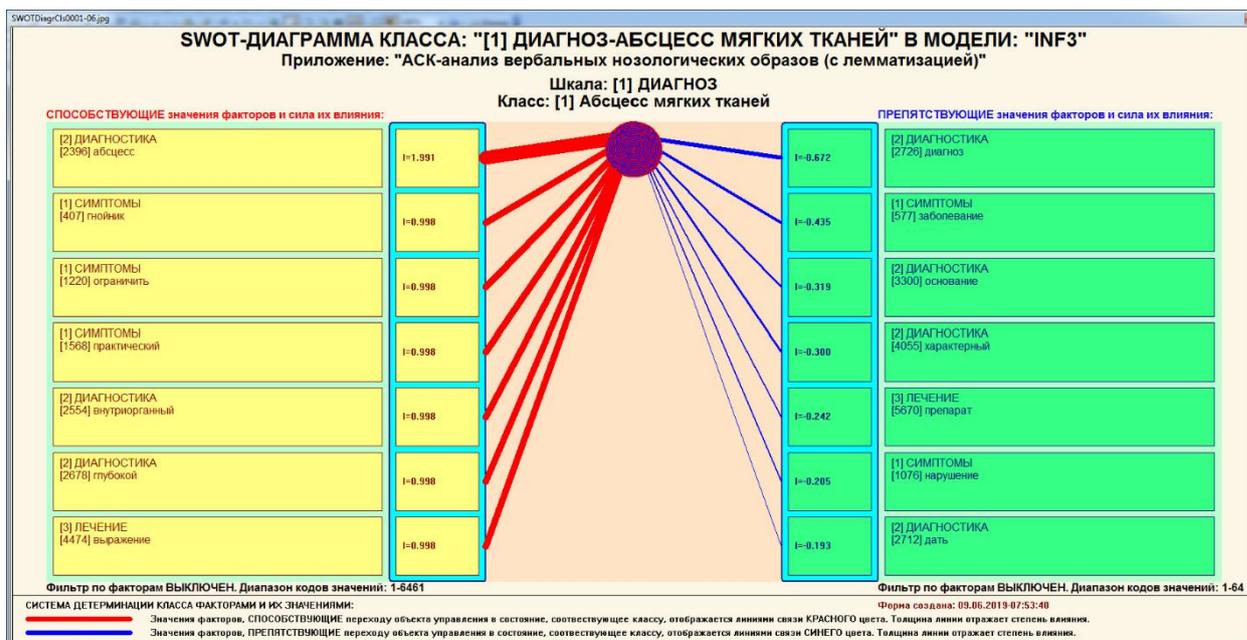


Рисунок 23 - Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (без фильтра) – SWOT-диаграмма

Также можно получить аналогичное семантическое ядро, но с применением фильтра по шкале «Симптомы». Информацию об этом мы можем получить из форм SWOT-анализа. Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» с включенным фильтром по шкале «Симптомы» изображены на рисунке 24.

То же самое проделываем, но уже с фильтром по шкале «Диагностика» (рисунок 25). Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (включен фильтр по шкале «Диагностика»)

Профессором Е.В. Луценко предложены нелокальные нейроны и нелокальные нейронные сети прямого счета, основанные на теории информации. Эти сети, предложенные автором 15 лет очень сходны с популярным сегодня нейробайесовским подходом.

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

**Выбор класса, соответствующего будущему состоянию объекта управления**

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	ДИАГНОЗ-Абсцесс мягких тканей	17,2282385	56	0,0000000
2	ДИАГНОЗ-Абсцесс носовой перегородки	12,6229590	50	0,0000000
3	ДИАГНОЗ-Аденоидные разрастания (аденоиды)	11,6046889	45	0,0000000
4	ДИАГНОЗ-Аденома простаты	17,0565378	64	0,0000000
5	ДИАГНОЗ-Адреногенитальные расстройства	20,5206774	54	0,0000000

**SWOT-анализ класса: 1 "ДИАГНОЗ-Абсцесс мягких тканей" в модели: 6 "INF3"**

**Способствующие факторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
407	СИМПТОМЫ-гнойник	0.998
1220	СИМПТОМЫ-ограничить	0.998
1568	СИМПТОМЫ-практический	0.998
913	СИМПТОМЫ-люба	0.995
762	СИМПТОМЫ-кишечные	0.993
9	СИМПТОМЫ-абсцесс	0.988
914	СИМПТОМЫ-любой	0.988
903	СИМПТОМЫ-локализация	0.986
2021	СИМПТОМЫ-стрептококк	0.974
1559	СИМПТОМЫ-позтому	0.972
1345	СИМПТОМЫ-палочка	0.970
2007	СИМПТОМЫ-стафилококк	0.970

**Препятствующие факторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
577	СИМПТОМЫ-заболевание	-0.435
1076	СИМПТОМЫ-нарушение	-0.205
1866	СИМПТОМЫ-симптом	-0.177
1648	СИМПТОМЫ-происходить	-0.153
1629	СИМПТОМЫ-причина	-0.151
1709	СИМПТОМЫ-развитие	-0.140
2031	СИМПТОМЫ-суставах	-0.135
819	СИМПТОМЫ-которое	-0.133
816	СИМПТОМЫ-котор	-0.130
815	СИМПТОМЫ-кость	-0.130
2265	СИМПТОМЫ-хронический	-0.121
1284	СИМПТОМЫ-острый	-0.114

ВКЛЮЧИТЬ фильтр по фактору    ВЫКЛЮЧИТЬ фильтр по фактору

Помощь    Abs    Prc1    Prc2    Inf1    Inf2    Inf3    Inf4    Inf5    Inf6    Inf7    SWOT-диаграмма

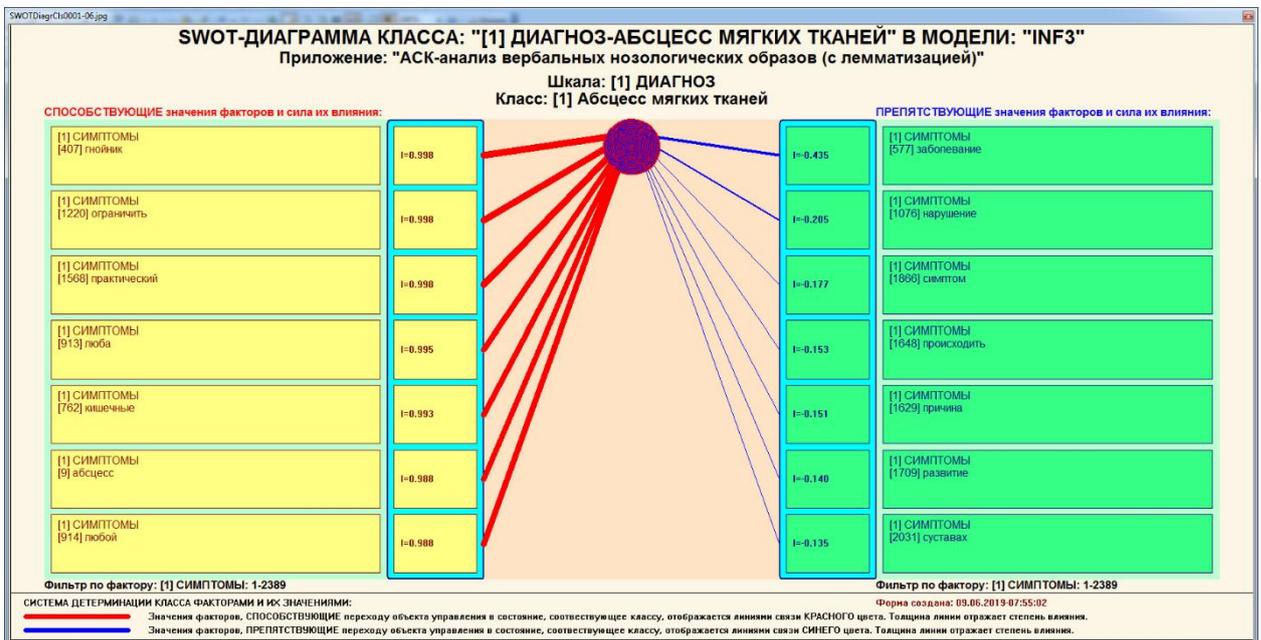


Рисунок 24 – Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (включен фильтр по шкале «Симптомы»)

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

**Выбор класса, соответствующего будущему состоянию объекта управления**

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	ДИАГНОЗ-Абсцесс мягких тканей	17,2282385	56	0,0000000
2	ДИАГНОЗ-Абсцесс носовой перегородки	12,6229590	50	0,0000000
3	ДИАГНОЗ-Аденоидные разрастания (аденоиды)	11,6046689	45	0,0000000
4	ДИАГНОЗ-Аденома простаты	17,0565378	64	0,0000000
5	ДИАГНОЗ-Адреногенитальные расстройства	20,5206774	54	0,0000000

**SWOT-анализ класса: 1 "ДИАГНОЗ-Абсцесс мягких тканей" в модели: 6 "INF3"**

**Способствующие факторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
2396	ДИАГНОСТИКА-абсцесс	1.991
2554	ДИАГНОСТИКА-внутриорганный	0.998
2678	ДИАГНОСТИКА-глубокой	0.998
3929	ДИАГНОСТИКА-томограф	0.995
3518	ДИАГНОСТИКА-потребоваться	0.991
3047	ДИАГНОСТИКА-локализация	0.988
3785	ДИАГНОСТИКА-сложность	0.984
3544	ДИАГНОСТИКА-представляем	0.970
2969	ДИАГНОСТИКА-компьютерный	0.963
4160	ДИАГНОСТИКА-явный	0.958
3988	ДИАГНОСТИКА-ультразвуковой	0.944
3580	ДИАГНОСТИКА-проведение	0.935

**Препятствующие факторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
2726	ДИАГНОСТИКА-диагноз	-0.672
3300	ДИАГНОСТИКА-основание	-0.319
4055	ДИАГНОСТИКА-характерный	-0.300
2712	ДИАГНОСТИКА-дать	-0.193
2786	ДИАГНОСТИКА-заболевание	-0.188
2727	ДИАГНОСТИКА-диагностик	-0.184
3586	ДИАГНОСТИКА-проводить	-0.177
3899	ДИАГНОСТИКА-также	-0.158
3158	ДИАГНОСТИКА-наличие	-0.158
3007	ДИАГНОСТИКА-кровь	-0.158
2908	ДИАГНОСТИКА-картины	-0.158
3852	ДИАГНОСТИКА-ставлюсь	-0.149

ВКЛЮЧИТЬ фильтр по фактору    ВЫКЛЮЧИТЬ фильтр по фактору

Помощь    Abs    Prc1    Prc2    Inf1    Inf2    Inf3    Inf4    Inf5    Inf6    Inf7    SWOT-диаграмма

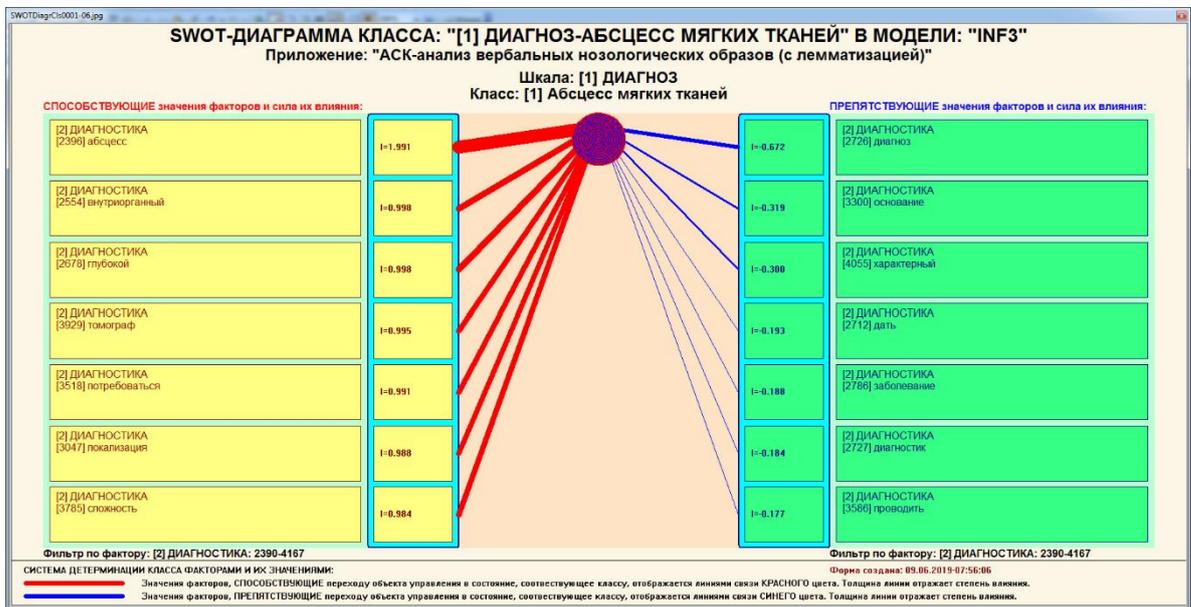


Рисунок 25 – Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (включен фильтр по шкале «Диагностика»)

Последняя интересующая нас информация, связанная с этим типом выходящих форм, это семантическое ядро с фильтром по шкале «Лечение».

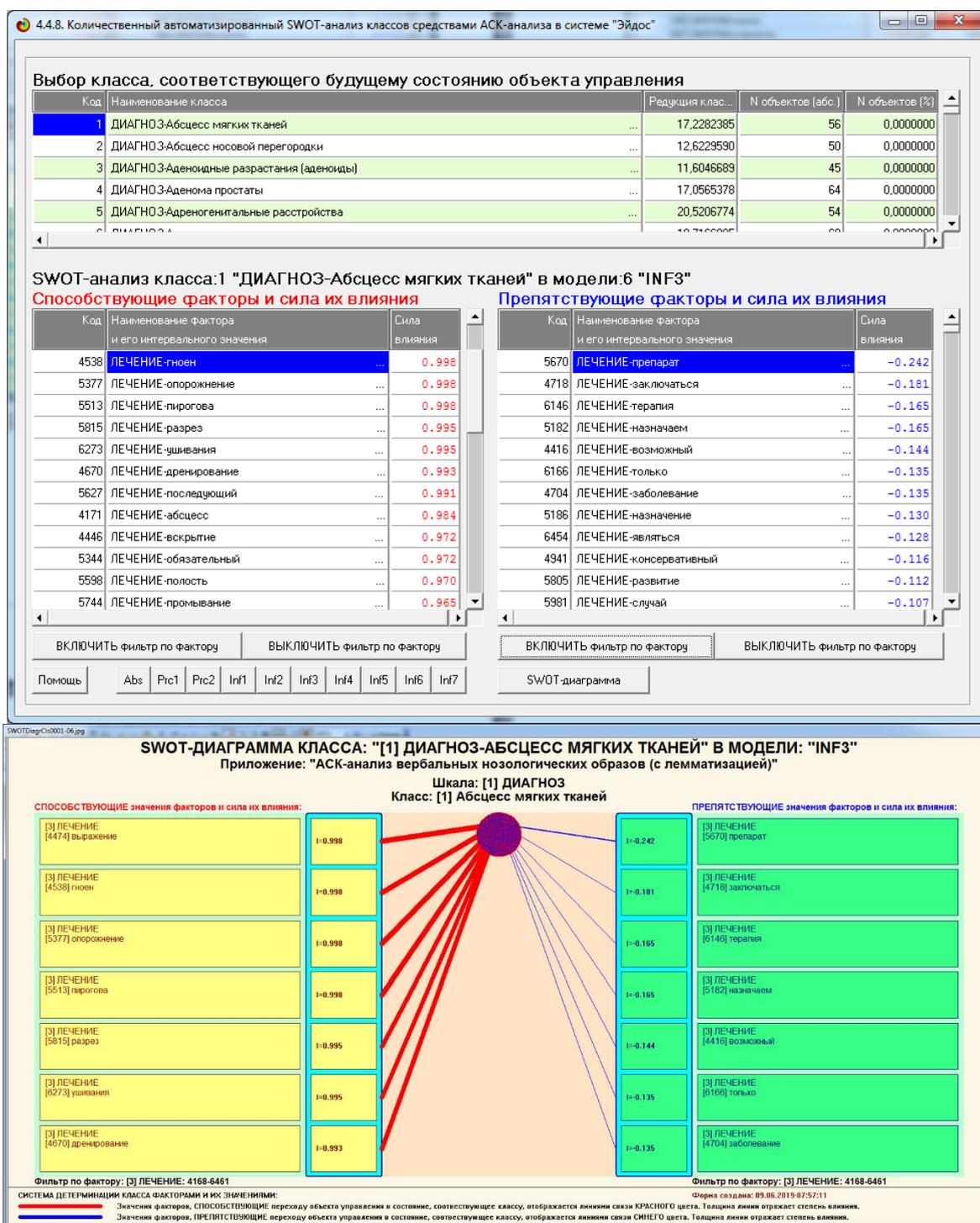


Рисунок 26 – Слова, наиболее характерные и нехарактерные для вербального нозологического образа «Абсцесс мягких тканей» (включен фильтр по шкале «Лечение»)

Пример нелокального нейрона группы абсцесс мягких тканей в СК-модели INF3 с 36 рецепторами приведен на рисунке 27.

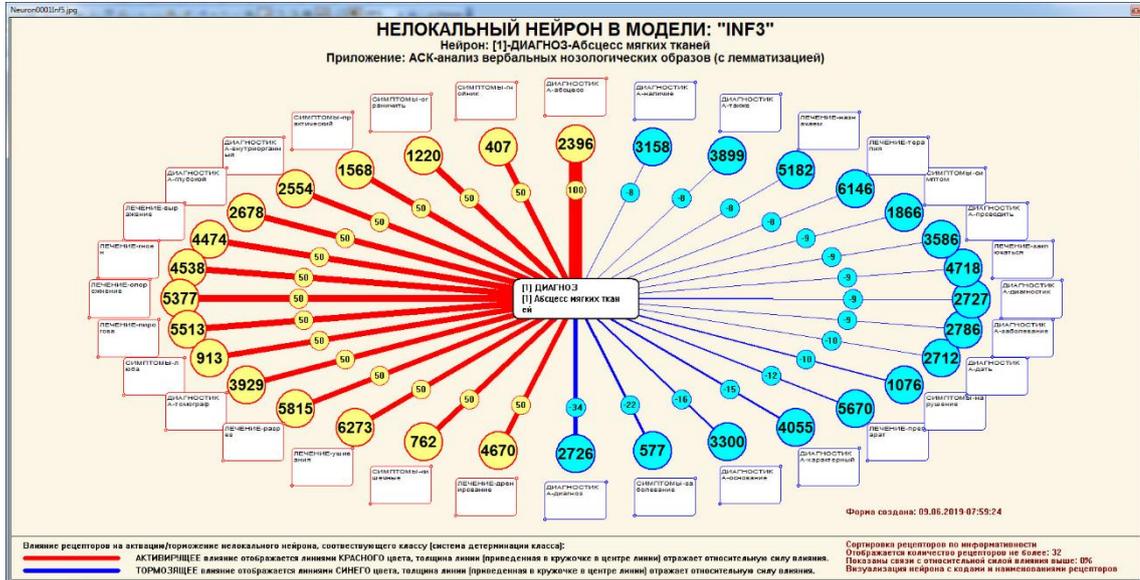


Рисунок 27 – Нелокальный нейрон «Абсцесс мягких тканей» 36 рецепторов

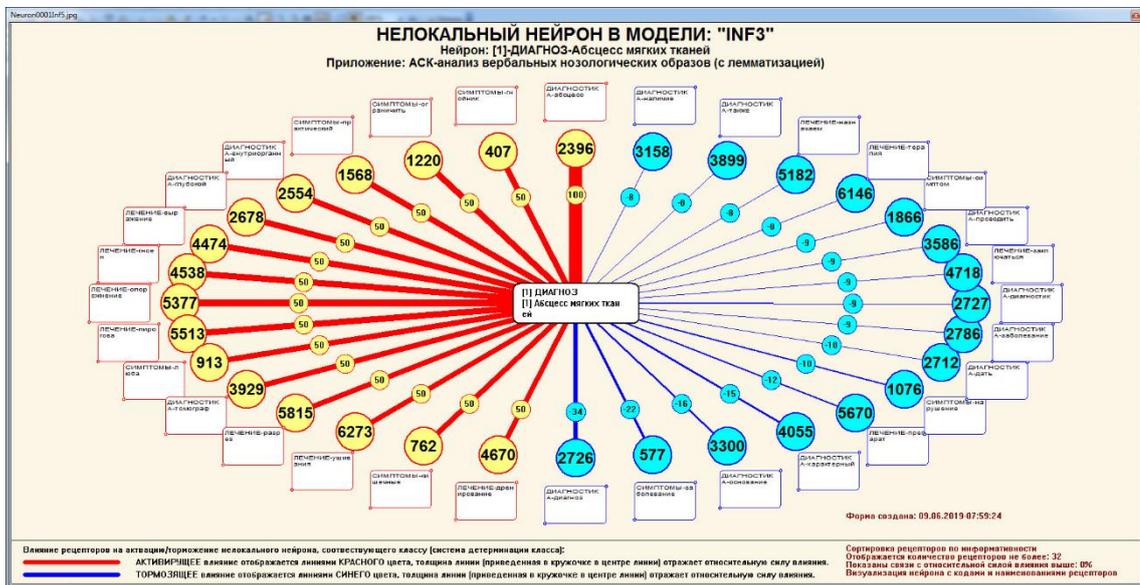


Рисунок 28. – Экранная форма визуализации парето-подмножества одного слоя нелокальной нейронной сети (показано 0,13% наиболее значимых синаптических связей)

Отметим, что системно-когнитивные модели (СК-модели), созданные в системе «Эйдос», отражают смысл каждого слова, который в данном случае представляет собой количество информации (в битах), содержащееся в слове, о принадлежности текста с данным словом к тому или иному классу (рисунок 29).

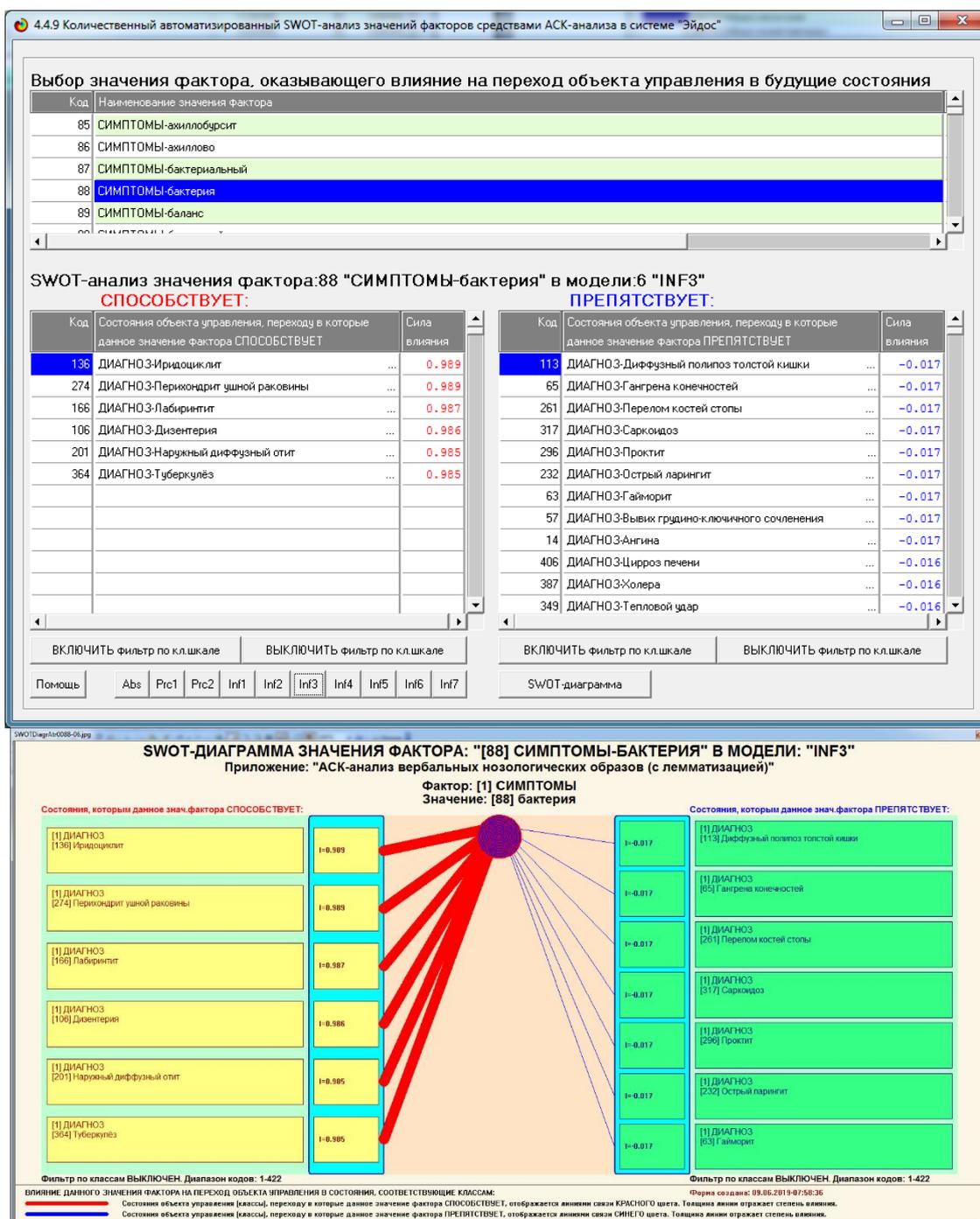


Рисунок 29 – Смысл симптома «бактерия»



4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса:1 "ДИАГНОЗ-Абсцесс мягких тканей" в модели:6 "INF3"

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	ДИАГНОЗ-Абсцесс мягких тканей	1	1	ДИАГНОЗ-Абсцесс мягких тканей	100.000
2	ДИАГНОЗ-Абсцесс носовой перегородки	2	120	ДИАГНОЗ-Заглочный абсцесс	21.498
3	ДИАГНОЗ-Аденоидные разрастания (аденоиды)	3	248	ДИАГНОЗ-Паратонзиллярный абсцесс	20.722
4	ДИАГНОЗ-Аденома простаты	4	2	ДИАГНОЗ-Абсцесс носовой перегородки	16.887
5	ДИАГНОЗ-Адреногенитальные расстройства	5	379	ДИАГНОЗ-Флегмона мягких тканей	15.686
6	ДИАГНОЗ-Акромегалия	6	242	ДИАГНОЗ-Панариций	15.384
7	ДИАГНОЗ-Алкогольная полинейропатия	7	276	ДИАГНОЗ-Пиодермия	14.708
8	ДИАГНОЗ-Алкогольный гепатит	8	161	ДИАГНОЗ-Костный панариций	12.916
9	ДИАГНОЗ-Аллергия	9	172	ДИАГНОЗ-Лимфаденит	11.915
10	ДИАГНОЗ-Алопеция	10	239	ДИАГНОЗ-Отогематома	11.853
11	ДИАГНОЗ-Альдостерома	11	303	ДИАГНОЗ-Разлитой перитонит	10.503
12	ДИАГНОЗ-Амиотрофический боковой склероз	12	383	ДИАГНОЗ-Фурункул наружного слухового прохода	10.325
13	ДИАГНОЗ-Анальная трещина	13	253	ДИАГНОЗ-Пельвиоперитонит	9.772
14	ДИАГНОЗ-Ангина	14	223	ДИАГНОЗ-Орхозпидидимит	9.691
15	ДИАГНОЗ-Ангина Симановского-Венсана (язв.)	15	384	ДИАГНОЗ-Фурункул носа	9.265
16	ДИАГНОЗ-Аневризма брюшной аорты	16	341	ДИАГНОЗ-Спонтанный пневмоторакс	9.052
17	ДИАГНОЗ-Анемия	17	143	ДИАГНОЗ-Карбункул	8.930
18	ДИАГНОЗ-Анкилоз коленного сустава	18	236	ДИАГНОЗ-Острый салпингит и оофорит	8.332
19	ДИАГНОЗ-Анкилоз локтевого сустава	19	369	ДИАГНОЗ-Урогенитальный хламидиоз	8.006
20	ДИАГНОЗ-Аномалии развития наружного и сре...	20	27	ДИАГНОЗ-Бедренная грыжа	7.486
21	ДИАГНОЗ-Аппендицит	21	235	ДИАГНОЗ-Острый ринит	7.331
22	ДИАГНОЗ-Аскаридоз	22	346	ДИАГНОЗ-Стрикура мочеочника	7.126
23	ДИАГНОЗ-Астигматизм	421	67	ДИАГНОЗ-Гастроэзофагеальная язва (пострезекционная)	-7.573

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ.фильтр по кл.шкале ВЫКЛ.фильтр по кл.шкале Вписать в окно Показать ВСЕ

Рисунок 31 – Результаты кластерно-конструктивного анализа классов

Тем не менее на рисунке 31 видно, что различие в распознавании диагноза абсцисс мягких тканей от следующей похожей болезни довольно велико, так что можно предположить, что определение этого диагноза будет довольно точным.

Несмотря на схожие диагнозы, путем пополнения базы программы специалисты будут постоянно ее совершенствовать. Программа будет выискивать особые отличия между схожими диагнозами и работать с каждым переобучение все точнее и точнее.

### 3.3.Описание алгоритма работы системы

Пациент приходит к специалисту, обычно общего назначения, к терапевту, и рассказывает ему про то, что его беспокоит, про симптомы его болезни.

Специалист вводит его жалобы в файл Excel с заранее подготовленной структурой. Первая строка должна содержать заголовки столбцов в следующем порядке: ФИО, симптомы. Далее занести этот файл в программу Эйдос, которая на основе разработанных верификационных моделей предоставит информацию относительно пациента. Также возможно составление файла на множество пациентов, и, допустим, в конце дня подводить итогов касательно всех пришедших пациентов сразу.

По каждому пациенту будет дан результат касательно его диагноза. Программа исходя из загруженной модели определения и файла симптомов предлагает наиболее достоверные результаты.

Специалист исходя из прогнозируемых диагнозов системы и своих профессиональных навыков может перенаправить пациента к более узко специализированному врачу, назначить дополнительное обследование, собрать консилиум врачей либо сам поставить диагноз. Специалист сам выбирает опираться ли ему на прогноз программы или нет. В тестовый период результаты могут быть не достаточно точными, но они уже с высокой вероятностью направят специалиста в правильную сторону, тем самым упростив его работу.

После поставленного диагноза, специалист заносит результаты обследования и правильный диагноз вместе с жалобами пациента в базу данных, на данных которой программа через определенное время будет переобучена, тем самым в программа с новой версией обучения будет определять диагнозы намного лучше. В перспективе обучить программу можно до такого уровня, что она сможет ставить диагнозы на уровне, а скорее всего лучше, чем специалисты в этой области.

### 3.4. Выводы

Создана система, которая позволяет прогнозировать вероятный диагноз пациента исходя из его симптомов. Система позволяет совершенствовать ее, путем предоставления ей правильных диагнозов со связанными с ними симптомами. Иными словами чем дольше будет существовать система, тем точнее будут ее результаты.

В перспективе после этапа тестирования можно создать подсистему для уже имеющийся системы ведения электронных карточек пациентов, ну или самостоятельную систему для сбора этих данных. А точнее симптомов и идентификационных данных пациентов, для более удобного взаимодействия с системой прогнозирования.

Система будет внедрена в тестовом режиме, для установления возможных ошибок и недостатков. Следующей стадией будет полноценное внедрение системы, и постоянное расширение ее базы знаний, что будет делать ее выводы точнее и точнее. Рассмотрим организационные и финансово-экономические аспекты внедрения собственной системы.

## 4. ОЦЕНКА ЭФФЕКТИВНОСТИ РАЗРАБОТАНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

### 4.1. План внедрения

Для внедрения системы необходимо установить на компьютер специалиста программный пакет Windows Office в котором присутствует необходимая программа Excel для редактирования таблиц в формате xls, или аналогичную программу. Например можно пользоваться бесплатной альтернативой LibreOffice или OpenOffice, которые поддерживает те же функции, но в основном используются на операционных системах Linux.

Далее потребуется скачать саму программу прогнозирования. Сделать это можно на официальном сайте профессора Луценко. Этот сайт на первом месте в поиске по аналогичному запросу. После установки системы, в нее необходимо загрузить сгенерированные базы, в которых хранятся наши обученные модели.

Необходимо подготовить файл в формате xls для записи в него симптомов пациентов. Его структура должна быть аналогична структуре того файла, на котором мы производили обучение, только в этом случае тот файл следует предварительно отчистить.

Также необходимо создать комиссию, которая будет проводить тестирование внедряемой системы. Смотреть на ее прогнозируемые данные, поправлять ее результаты, вносить исправленные данные обратно в систему, чтобы она могла совершенствоваться и в следующий раз выдавать результаты более достоверные.

После создания комиссии необходимо выполнить непосредственно тестирование. Так как проект представляет из себя систему, то логичными будут для него функциональное тестирование.

После проведения тестирования были установлены несоответствия, проведена доработка и выполнено повторное тестирование.

Далее проходит создание комиссии по внедрению. После проводится внедрение в рабочий процесс, ответственный сотрудник изучил особенности системы и было проведено внедрение.

#### 4.2. Оценка эффективности

В медицинском учреждении основным методом постановки диагнозов являлся личный опыт и профессиональные качества работающего специалиста, выявляющего закономерности исключительно полагаясь на свои собственные знания медицинских болезней. Метод неэффективен в силу мощнейшего влияния человеческого фактора. К тому же компания была зависима от опытного медицинского персонала и была не способна в случае потери квалифицированного кадра в сжатые сроки подготовить новые кадры.

Новая система прогнозирования диагнозов позволяет помогать специалисту ставить диагнозы. В легких случаях определения болезни система приносит не такую большую пользу, зато если попадаете пациент со сложным случаем, система во много раз повышает скорость процесса постановки диагноза. Она позволяет сразу же определить наиболее подходящие диагнозы для данного случая, даже если не может определить точный. Специалисту становится гораздо проще определиться, имея перед собой небольшой список возможных вариантов. В некоторых случаях все варианты указывают на диагнозы, которые может выносить только другой узкоспециализированный специалист. Тогда терапевт врач общей диагностики точно знает к какому специалисту направить пациента. Без аналогичной системы, как система Эйдос, врачи часто ошибались с направлением пациентов, что приводило к лишним обследованиям и трате времени как самого пациента, так и других врачей, обследовавших его зря. Новая система экономит время пациента, те-

рапевта и остальных узкоспециализированных врачей. Система внедрена на предприятии в тестовом режиме 10.04.2019.

После внедрения увеличился приток клиентов, повышена скорость и качество поставленных диагнозов. Сами решения стали более адекватными. Удалось сократить чистое рабочее время специалистов, 30% времени которых занимали обдумывания и чтения литературы касаясь пациентов. В появившееся время врачи занимаются более научной сферой медицины, разрабатывая и изучая что-то новое. Уменьшилось время приема пациента примерно на 60%, что с учетом увеличенного отдыха специалиста позволяет принимать примерно на 30% больше пациентов в день. Благодаря увеличению времени, которое тратится на научные исследования, специалисты, тестирующие данную систему стали больше публиковать статей и научных работ в области медицины. В общем специалисты стали менее загружены, а вследствие более жизнерадостней. Их отношение к принимаемым пациентам улучшилось. Пациенты, с которыми теперь врачи обходятся более доброжелательно, от хорошего отношения чувствуют себя лучше и меньше жалуются на несуществующие недуги.

#### 4.3. Определение плановой себестоимости проведения работ

Несмотря на хорошую эффективность системы её разработка не включает в себя большие траты.

Аналитическая система, составляя семантические ядра по сути решает задачу области SEO. В случае разработки нового электронного справочника болезней, можно воспользоваться семантическими ядрами болезней, построенными с помощью этой системы, для оптимизации выдачи в поисковых запросах.

Поисковые системы умны, но им по-прежнему нужна помощь. Основные двигатели всегда работают над улучшением своих технологий, чтобы

глубже проникнуть в Интернет и получить лучшие результаты для пользователей. Тем не менее, существует предел тому, как искать поисковые системы. В результатах поиска вы можете увидеть качество результатов SEO.

Помимо того, что контент доступен для поисковых систем, SEO также помогает повысить рейтинг. Интернет становится все более конкурентоспособным и тем, кто выполняет SEO.

Поисковая оптимизация (SEO) в 2019 году - это технический, аналитический и творческий процесс для улучшения видимости вебсайта в поисковых системах. Основная функция SEO - привлечь больше неоплаченного полезного трафика на сайт, который преобразуется в продажи.

Бесплатные советы по SEO, которые вы хотите прочитать на этой странице, помогут вам создать успешный вебсайт с поддержкой SEO.

Для получения данных реальной эффективности системы необходимо время. За продолжительный период работы системы необходимо сравнить количеств принятых правильно диагнозов и неправильно, а также скорость вынесения их медицинского учреждения при традиционном подходе постановки диагнозов и при использовании интеллектуальной системы принятия решений. Необходимо спрашивать сотрудников, какое бы решение они приняли, если бы они работали по старой модели. Необходимо проведение полно лечения и последующей повторной проверки самочувствия больного. После чего сравнить их результат с результатом, полученным с использованием интеллектуальной системы управления. На основе этого необходимо будет сделать окончательный вывод об эффективности системы.

На данном этапе оценка и реальное определение эффективности предполагает анализ последствий путем сравнения с традиционным подходом. Но в силу ограниченности по времени можно только сделать предварительные выводы о эффективности работы системы.

Затраты на систему Эйдос в расчет не берём, так как это открытое программное обеспечение с бесплатным открытым исходным кодом. В связи с чем затрат на её использование фирма не несёт.

Прямой экономической выгоды данная система не приносит, что не так уж и важно для бюджетного государственного учреждения. Но система существенно экономит время примерно 60% с каждого приема, за счет чего удалось увеличить количество принимаемых людей в день на 30% и увеличить свободное время врача, а также увеличить время научных исследовательских работ. В общем случае отзывы о работе специалистов проводящих тестирование данной системы стали лучше.

#### 4.4. Вывод

Создание и внедрение системы было целесообразно со всех точек зрения. Поставленная задача автоматизации работы персонала, сокращения расходов и издержек успешно решена. Проблемы неэффективной траты рабочего времени, отсутствия систематизации и централизованного анализа были успешно преодолены.

## ЗАКЛЮЧЕНИЕ

В процессе выполнения поставленной задачи была произведена оценка и составлена характеристика объекта исследования – интерактивного интеллектуального справочника болезней внедренного в тестовом режиме в Краснодарскую краевую больницу №1 им. Очаповского. Рассмотрена организационная структура и фактические методы функционирования. Рассмотрен процесс постановки диагноза пациенту и назначения ему лечения.

Выявлены и структурированы принципы работы медицинского учреждения, из чего в последствии создана модель разрабатываемой системы.

Выявлены недостатки и места подверженные автоматизированию. Произведена постановка и обозначение задач, решение которых позволит увеличить скорость а так же точность постановки диагнозов пациентам. Выделена главная задача – определение наиболее достоверного диагноза пациента.

Выявлены внешние и внутренние информационные связи проектируемой системы. Описана построенная на их основе структурная схема объекта, с обоснованием выбора основных функциональных блоков. Обоснованы и описаны схемы алгоритма функционирования объекта.

Обоснована необходимость, целесообразность и экономическая эффективность разрабатываемой системы.

Произведено внедрение системы в медицинском учреждении как тестового варианта в целях тестирования основного функционала и совершенствования системы.

## СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. Барановская Т.П. Методики и модели разработки архитектуры предприятия: Монография / Т.П. Барановская, А.Е. Вострокнутов – Краснодар, 2016. – 205 с.
2. Васильев А. Программирование на С# для начинающих. Основные сведения. / Алексей Васильев – Эксмо, 2018. – 592 с.
3. Википедия [Электронный ресурс]. URL: <https://ru.wikipedia.org>
4. Джеймс К. Компьютерные сети / Куроуз Джеймс, Росс Кит – Эксмо, 2016. – 912 с.
5. Каталог API (Microsoft) и справочных материалов [Электронный ресурс]. URL: <https://msdn.microsoft.com/library>.
6. КиберФорум - форум программистов, системных администраторов и администраторов баз данных [Электронный ресурс]. URL: <http://www.Cyberfo-rum.ru/>
7. Лойко В.И. Методические рекомендации по подготовке к государственной итоговой аттестации. / В.И. Лойко и др.– Краснодар: КубГАУ, 2019. – 75 с.
8. CodeNet - все для программиста [Электронный ресурс]. URL: <http://www.codenet.ru/>
9. Stack Overflow – сайт вопросов и ответов для программистов. [Электронный ресурс]. URL: <https://ru.stackoverflow.com/>
10. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2017. – 605 с.  
<http://elibrary.ru/item.asp?id=18632909>

11. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с. ISBN 978-5-94672-757-0. <http://elibrary.ru/item.asp?id=21358220>

12. Луценко Е.В. Применение АСК-анализа и интеллектуальной системы "Эйдос" для решения в общем виде задачи идентификации литературных источников и авторов по стандартным, нестандартным и некорректным библиографическим описаниям / Е.В.Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №09(103). С. 498 – 544. – IDA [article ID]:1031409032. – Режим доступа: <http://ej.kubagro.ru/2014/09/pdf/32.pdf>, 2,938 у.п.л.

13. Луценко Е.В., Открытая масштабируемая интерактивная интеллектуальная online среда «Эйдос» («Эйдос-online»). Свид. РосПатента РФ на программу для ЭВМ, Заявка № 2017618053 от 07.08.2017, Гос.рег.№ 2017661153, зарегистр. 04.10.2017. – Режим доступа: <http://lc.kubagro.ru/aidos/2017661153.jpg>, 2 у.п.л.

14. Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализеи системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №02(126). С. 1 – 32. – IDA [article ID]: 1261702001. – Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 у.п.л.

15. Кириллов, В.В. «Введение в реляционные базы данных. Введение в реляционные базы данных» / В.В. Кириллов, Г.Ю. Громов. - СПб.: БХВ-Петербург, 2018. - 464 с.

16. Саммерфилд, М. «Программирование на Python 3» - СПб.: Символ-Плюс, 2017год. - 390 С.

## ПРИЛОЖЕНИЕ 1. ЛИСТИНГ ПРОГРАММЫ

Исходный текст программы Parse.py парсинга сайта-базы болезней  
medaboutme.ru

```
# pip install requests
# pip install pandas
# pip install openpyxl
# pip install beautifulsoup4
# pip install lxml

import requests
from bs4 import BeautifulSoup
import csv
from datetime import datetime
from multiprocessing import Pool
import re

from openpyxl import Workbook
from openpyxl.utils.dataframe import dataframe_to_rows
import pandas as pd

MAIN_URL = "https://medaboutme.ru"
PARSE_URL = "https://medaboutme.ru/zdorove/spravochnik/bolezni/"
patternCol3 = re.compile("симптомы")
patternCol4 = re.compile("диагностика")
patternCol5 = re.compile("лечение")

filename = "excelFile.xlsx"

counterError = 0
```

```
counterRows = 0
```

```
def get_html(url):
```

```
    r = requests.get(url)
```

```
    return r.text
```

```
def get_all_links(html):
```

```
    soup = BeautifulSoup(html, 'lxml')
```

```
    lies = soup.find('div', class_="glossary-letters").find_all("li")
```

```
    links = []
```

```
    count = 0
```

```
    for li in lies:
```

```
        count+=1
```

```
        a = li.find('a').get("href")
```

```
        link = MAIN_URL + a
```

```
        links.append(link)
```

```
    print("Количество ссылок " + str(count))
```

```
    return links
```

```
def get_page_data(html):
```

```
    global counterRows
```

```
    global counterError
```

```
    soup = BeautifulSoup(html, 'lxml')
```

```
    try:
```

```
        name = soup.find('h1', class_="width-comp").text.strip()
```

except:

```
name = "не найдено"
```

```
cols = soup.find_all('div', class_="disease-detail-item")
```

```
col3 = "не найдено"#симптомы
```

```
col4 = "не найдено"#диагностика
```

```
col5 = "не найдено"#лечение
```

```
for col in cols:
```

```
    try:
```

```
        colName = col.find('div', class_="mobile disease-detail-  
header").text.strip().lower()
```

```
        # print('find '+ colName)
```

```
        if (bool(patternCol3.search(colName))):#симптомы
```

```
            try:
```

```
                col3 = col.find('div', class_="disease-detail-body").find('p').text
```

```
            except:
```

```
                col3 = "не найдено"
```

```
                counterError += 1
```

```
        if (bool(patternCol4.search(colName))):#диагностика
```

```
            try:
```

```
                col4 = col.find('div', class_="disease-detail-body").find('p').text
```

```
            except:
```

```
                col4 = "не найдено"
```

```
                counterError += 1
```

```
        if (bool(patternCol5.search(colName))):#лечение
```

```
            try:
```

```
                col5 = col.find('div', class_="disease-detail-body").find('p').text
```

```
except:
    col5 = "не найдено"
    counterError += 1
```

```
except:
    pass
```

```
data = {'name' : name,
        'col3' : col3,
        'col4' : col4,
        'col5' : col5}
# print(data)
counterRows += 1
print("row: " + str(counterRows) + " errors: " + str(counterError))
return data
```

```
def print_data_row(data):
    print('_____')
    print(data['name'])
    print(data['col3'])
    print(data['col4'])
    print(data['col5'])
    print('_____')
```

```
def mass_to_data_frame(massData):
    d = {'диагноз' : [],
        'диагноз2' : [],
        'симптомы' : [],
        'диагностика': [],
        'лечение' : []}
```

```
for row in massData:
    d['диагноз'].append(row['name'])
    d['диагноз2'].append(row['name'])
    d['симптомы'].append(row['col3'])
    d['диагностика'].append(row['col4'])
    d['лечение'].append(row['col5'])
```

```
df = pd.DataFrame(d)
print("dataFrame")
print(df)
return(df)
```

```
def write_excel(dataFrame):
    wb = Workbook()
    # wb.save(filename)
    sheet1 = wb.create_sheet("Лист1", 0)
    print(sheet1)
    # wb.create_sheet("Лист2", 1)
    # wb.create_sheet("Лист3", 2)
    # sheet1 = wb['Лист1']
    # sheet1 = wb.active

    for x in dataframe_to_rows(dataFrame):
        # print("row")
        # print(x)
        sheet1.append(x)

    wb.save(filename)
```

```

# wb.save(filename)

def make_all(all_links):
    massData = []

    for i,url in enumerate(all_links):
        data = get_page_data(get_html(url))
        massData.append(data)
        # print_data_row(data)
        # if i == 1: break

    dataFrame = mass_to_data_frame(massData)
    write_excel(dataFrame)

def main():
    start = datetime.now()
    all_links = get_all_links(get_html(PARSE_URL))
    make_all(all_links)

    end = datetime.now()
    print(str(end - start))

if __name__ == '__main__':
    main()

```