

Санкт-Петербургский государственный университет

Экономический факультет



ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
по направлению «Бизнес-информатика»

**РАЗРАБОТКА МОДУЛЯ ДЛЯ ПЕРВИЧНОЙ САМОДИАГНОСТИКИ
ПАЦИЕНТОВ, ОСНОВАННОГО НА НЕЙРОННЫХ СЕТЯХ,
В РАМКАХ РАЗВИТИЯ ТЕЛЕМЕДИЦИНЫ**

Выполнила:

студентка 4 курса

Шмагина Вероника Александровна

_____ / _____

Научный руководитель:

доктор технических наук, доцент

Лебедев Илья Сергеевич

_____ / _____

Санкт-Петербург

2018 г

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
ГЛАВА 1. ПОНЯТИЕ ТЕЛЕМЕДИЦИНЫ И ЕЕ РАЗВИТИЕ В РОССИИ.....	6
1.1. Понятие телемедицины	6
1.2. Здоровоохранение в России в рамках развития телемедицины	10
1.3. Использование нейронных сетей в условиях телемедицины	15
ГЛАВА 2. РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ПЕРВИЧНОЙ САМОДИАГНОСТИКИ ОСНОВАННОГО НА НЕЙРОННЫХ СЕТЯХ	21
2.1. Описание приложения	21
2.1.1. Принцип работы приложения.....	21
2.1.2. Построение информационно-логической модели (ИЛМ).....	22
2.2. Реализация нейросетевого модуля	23
2.3. Анализ качества модели	30
2.4. Оценка стоимости разработки программного обеспечения	33
2.5. Результаты внедрения приложения.....	35
ЗАКЛЮЧЕНИЕ.....	38
СПИСОК ЛИТЕРАТУРЫ	39

ВВЕДЕНИЕ

Актуальность

В современном мире все больший интерес проявляется в области изучения нейронных сетей и нейрокомпьютеров. Актуальность исследований в этом направлении подтверждается большим числом публикаций, конференций и различных практических применений в областях экономики, медицины, химии, физики и многих других. С помощью искусственных нейронных сетей можно выполнять, например, распознавание объектов на изображении, предсказывать колебания на рынке валют, моделировать модель принятия решений, определять направления научного поиска, синтезировать речь по тексту, ставить диагнозы и выявлять патологии. На западе применение искусственных нейронных сетей (ИНС) уже достаточно обширно, в России же это воспринимается как что-то экзотичное. Существует лишь небольшое количество российских фирм, которые используют ИНС в практических целях.

Спектр областей внедрения нейронных сетей достаточно широк в современном мире, в данной исследовательской работе была выделена одна из самых на сегодняшний день популярных сфер их применения - медицина. Актуальность выбора этой области также обусловлена и тем, что в августе 2017 года были внесены в отдельные законодательные акты Российской Федерации по вопросам применения информационно-телекоммуникационных технологий в сфере охраны здоровья граждан и создания национальных научно - медицинских центров, которые легитимизируют удаленные консультации и мониторинг пациентов.¹

ИНС может делать диагнозы с высокой точностью, а главное такую технологию можно использовать для определения первичного диагноза или наблюдения. Все это поднимет диагностику на новый уровень. При этом не требуется создавать что-то абсолютно новое, необходимо просто приспособить имеющиеся алгоритмы и железо. В данной дипломной работе предлагается использование нейронных сетей для постановки первичного самодиагноза.

Необходимость создания в России приложения на основе нейронных сетей также вызвано и тем, что поликлиники в большинстве городов не справляются с потоком пациентов, тем самым не оказывают своевременную помощь. Таким образом, опасность болезни

¹ Федеральный закон от 21.11.2011 (ред. 15.06.2017) N 323-ФЗ "Об основах охраны здоровья граждан в Российской Федерации".

увеличиваются и на ликвидацию ее затрачивается больше денежных средств, как со стороны пациента, так и государства.

Объект исследования

Объектом исследования данной дипломной работы является определение эффективности внедрения нейронных сетей в сферу здравоохранения России.

Предмет исследования

Предметом исследования является эффект от внедрения приложения для первичной самодиагностики на основе нейронных сетей в систему здравоохранения в рамках развития телемедицины.

Цель исследования

Целью данного исследования являлась разработка приложения, которое сможет увеличить скорость и точность постановки диагноза, а как следствие позволит увеличить уровень жизни в стране.

Для достижения данной цели были поставлены следующие **задачи**:

1. Изучить состояние системы здравоохранения в современной России;
2. Изучить использование нейронных сетей в сфере медицины;
3. Разработать принцип работы приложения для постановки первичного самодиагноза;
4. Реализовать нейросетевой модуль для анализа качества модели;
5. Оценить затраты на создание приложения и эффект от его разработки и запуска.

Структура исследования

Дипломная работа состоит из введения, двух глав, заключения и списка литературы. Общий объем работы 41 страница основного текста.

Первая глава – теоретическая, здесь описывается понятие телемедицины и дана краткая история ее появления и развития. Также в данной главе рассматривается использование телемедицинских сервисов в России, состояние системы здравоохранения на данный момент и предлагается идея внедрения нейронной сети для совершенствования здравоохранительной отрасли в России.

Вторая глава – практическая, в данной главе описан принцип работы приложения, описан разработанный модуль для постановки первичного самодиагноза, а также проведен анализ качества модели и сравнительный анализ системы здравоохранения на данный момент с системой при использовании нейронных сетей.

ГЛАВА 1. ПОНЯТИЕ ТЕЛЕМЕДИЦИНЫ И ЕЕ РАЗВИТИЕ В РОССИИ

1.1. Понятие телемедицины

В последние десятилетия прогресс в развитии компьютерных технологий приводит к качественным изменениям в жизни общества. Сейчас становятся очень популярными такие понятия как «цифровая экономика», «электронное правительство», «электронные услуги», «телемедицина» и так далее.

История развития телемедицины началась в начале XX века, когда появилась телефонная связь. Связь с доктором, находящемся на расстоянии, стала гораздо доступнее, однако качество таких консультаций ограничивалось качеством связи. С появлением телеграфа стало возможным передавать медицинские данные, а с помощью радио оказывать медицинскую помощь на судах морского флота. В связи с внедрением новых технологий, в течение нескольких лет в научных архивах появлялись соответствующая терминология. Например, в 1905 году Виллем Эйтховен впервые реализовал передачу электрокардиограммы по телефону, этот способ впоследствии был назван «телекардиографией». Для телеконсультаций с помощью радио применялся термин «телегнозия». А с появлением узких направлений стали появляться такие термины, как «телепсихиатрия», «теледиагностика», «телерадиология» и многие другие [9].

Такое массовое распространение сходных терминов для обозначения близких по сути вещей требовало систематизации и упорядочивания их использования. Понятие телемедицины появилось еще в 70-х годах XX века, и объединяло оно множество телекоммуникационных и информационных методов, применяемых в здравоохранении, а также их разнообразные клинические приложения. К 2007 Всемирная Организация Здравоохранения расширила значение данного термина. Телемедицина - предоставление медицинских услуг, когда расстояние является решающим фактором, всеми специалистами здравоохранения, использующими информационные и коммуникационные технологии для обмена достоверной информации для диагностики, лечения и профилактики болезней и травм, исследований и оценки, а также для непрерывного образования медицинских работников, все в интересах продвижения здоровья людей и их общин [29].

В современном мире уже невозможно представить свою жизнь без смартфона и компьютера. Технологическая революция помогла сделать техническое обеспечение телемедицины гораздо дешевле, чем это было раньше. Многие современные программные решения для телекоммуникаций требуют всего лишь компьютер или смартфон и интернет,

чтобы подключиться к программному обеспечению для виртуального общения с пациентом. Сетевые технологии предоставляют также возможность документальной передачи истории болезни при переводе больных из клиники в клинику, оперативное решение вопросов страхования и оплаты.

Многие страны и международные организации создают различные телемедицинские проекты. Во Всемирной организации здравоохранения рассматривается проект о создании глобальной сети телекоммуникаций в медицине, которая позволит ускоренно обмениваться научными документами и информацией, участвовать в видеоконференциях и заочных дискуссиях и совещаниях по всему миру.

Цель телемедицины – предоставление качественной медицинской помощи любому человеку, независимо от его местонахождения и социального положения [5]. Существуют четыре основных типа телемедицины:

- Дистанционный мониторинг

Дистанционный мониторинг – метод, который позволяет специалистам отслеживать жизненные показатели и активность пациентов на расстоянии. Этот тип мониторинга часто используется для людей, чье здоровье находится в высокой степени риска, например, с сердечными заболеваниями или люди, которые только выписались из больницы. Также данный метод очень полезен, когда речь идет о хронических заболеваниях. Это может быть использовано диабетиками, например, для измерения уровня глюкозы в крови и отправки данных своему лечащему врачу, который сможет регулировать лечение по необходимости.

- Телемедицина в режиме реального времени

Под телемедициной в режиме реального времени подразумевается видео-общение в режиме реального времени. В отличие от остальных типов телемедицины, данный тип не требует посещения врача для осмотра. Очень важным является тот факт, что при таком приеме неуместным является использование типичных социальных сетей или средств для конференций. Такой тип приема сопровождается специальным программным обеспечением, которое защищает конфиденциальность пациента и выполняет все меры безопасности.

- Телеобучение

Данный тип телемедицины заключается в участии в конференциях, видеосеминарах, лекциях с использованием телекоммуникационного оборудования. Врач во время таких мероприятий может повышать свой уровень профессионального образования без отрыва от места работы, например, данная технология может использоваться в режиме «теленаставничества» при проведении хирургических операций, когда более опытный специалист наблюдает и контролирует действия специалиста.

- Мобильные телемедицинские комплексы

Телемедицинский комплекс представляет собой совокупность различных мобильных и портативных программно-аппаратных средств, которые могут быть применены в домашней телемедицине, экстремальной и военной медицине, а также в скорой и амбулаторной помощи и реабилитации пациентов, например, тонометр, электрокардиограф, спирогроф и глюкометр.

Каждый из данных типов телемедицины имеет множество преимуществ как для пациентов, так и для врачей, что позволяет ей стремительно развиваться. Для врачей были выявлены следующие преимущества [32]:

- Позволяет расширить клиентскую базу специалиста.

Телемедицина предлагает врачам возможность охватить людей, которые иначе не имели бы доступа по разным причинам к пациентам, например, если пациент находится в сельской местности или находится в другой стране.

- Дополнительные потоки доходов

Увеличение базы пациентов означает для врачей одно – больше доходов. Кроме того, данный вид медицины дает шанс проводить прием в комфортном месте, например, дома. Также нет необходимости тратить время на дорогу, если речь идет о враче на выезде, что позволяет в освободившееся время провести дополнительный прием. Более того, врачи могут получать дополнительные премии и субсидии от государства с целью развития телемедицины.

- Снижение количества отмененных приемов

Телемедицина позволяет снизить количество аннулированных приемов из-за своего удобства. Так как пациенты могут находиться где угодно и не иметь несколько часов времени, чтобы добраться до специалиста в поликлинике, они отменяют или просто не приходят на прием к врачу. Благодаря дистанционному общению, пациент может из любого места связаться с врачом, потратив при этом значительно меньше времени.

- Возможность получать дополнительное образование от специалистов

Телемедицина является ресурсом для повышения квалификации врачей, средством проведения консилиумов и телеконференции. С помощью такого шанса врачи могут получать навыки и практики от лучших специалистов планеты.

Для пациентов существует также значительный ряд преимуществ, среди них такие, как:

- Отсутствие потери времени и денег на путь к врачу

Когда пациент проводит консультацию со своим доктором посредством мобильного телефона или компьютера, он экономит свои деньги на топливе, парковке для машины или на общественном транспорте. Более того, нет траты времени на дорогу или риска, что попадешь в пробку, что приведет к тому, что будет опоздание на прием или возвращение на работу.

- Нет необходимость брать отгул или прерывать работу ради приема у врача.

Для данного типа приема можно легко назначить встречу в течение обеденного перерыв, до или после работы. Можно находиться в любом месте, который предоставляет достаточный уровень конфиденциальности

- Устранение проблем связанных с транспортировкой людей, имеющих проблемы с двигательным аппаратом.

Благодаря телемедицине целый слой людей, которые не могут свободно перемещаться, но им необходимо получать регулярную медицинскую помощь, смогут это сделать, не прося помощи родных и близких и без больших сложностей для самих себя.

- Отсутствие необходимости поиска сиделки за престарелыми или няни для детей.

Многие люди имеют необходимость заботиться о детях или престарелых родителях. Поиск замены для того, чтобы увидеть доктора может быть дорогим и затруднительным. К счастью, телемедицина позволяет и увидеть доктора, и продолжать присматривать за близкими.

- Помощь в любой момент по необходимости.

Не всегда можно записаться на прием в ближайшей поликлинике или к своему врачу, как только появилась эта необходимость. Иногда требуется ждать несколько дней, а то и недель, чтобы с ним встретиться. Встретиться с онлайн-врачом можно по первой необходимости, потому что существует множество врачей различных специальностей, все из которых находятся «близко» и смогут проконсультировать по нужному вопросу.

- Доступ к узким специалистам, находящимся в удалении.

Некоторым пациентам, которым необходима помощь узкого специалиста приходится ехать на далекие расстояния и вкладывать много средств и времени на это. Телемедицина дает возможность пациенту и его лечащему врачу получить консультацию опытного специалиста, который находится на расстоянии. Когда речь идёт о серьезных проблемах необходима консультация лучших, а не ближайших.

- Снижение траты времени в процессе ожидания приема.

Часто бывает, что приходится в комнате ожидания ждать приглашение специалиста, читая газету или просто теряя время впустую. Телемедицина позволяет ожидать приема занимаясь работой или домашними делами. Среднее время ожидания у врача составляет 35,94 минут [31].

- Более здоровый организм.

Когда появляется возможность встречаться с доктором так часто как это нужно, без необходимости уходить с работы и тратить время, можно лучше управлять своим лечением, стилем жизни и хроническими проблемами, которые могут быть.

Подводя итоги всему вышесказанному, благодаря использованию информационных технологий в здравоохранении появились дополнительные возможности применения новых лечебных подходов, значительно расширяющих круг пациентов, которые находятся под амбулаторным наблюдением, и уменьшающих количество визитов пациентов в клинику. Например, по данным Американской телемедицинской ассоциации число очных встреч с врачами в Америке сократилось на 70% с 2008 по 2016 год. Также повышается удовлетворенность больных медицинскими услугами, повышается информированность пациентов о своих заболеваниях и улучшается качество обслуживания и своевременная коррекция лекарственной терапии, а, следовательно, снижается частота госпитализаций и обращений за скорой медицинской помощью. Конечно, состояние ее в разных частях мира отличается и по степени внедрения, и по полноте реализации, однако, цели и задачи, а также положительные моменты этой новой технологии идентичны, и ее необходимо развивать.

1.2. Здравоохранение в России в рамках развития телемедицины

История развития телемедицины в России начинается в 70-х годах XX века, когда медицинские сотрудники начали пересылать результаты ЭКГ в специализированные центры для получения консультаций. Следующим значительным шагом развития данного направления было проведение первой медицинской видеоконсультации врачами Российской военно-медицинской академии в 1995 году. А в 2001 году произошла интеграция российских телемедицинских сетей с мировой информационной системой, что стало новым глобальным шагом в развитии телемедицины. Первый закон по вопросам о телемедицине в России появился 20 декабря 2012 г., этот закон создал Координационный совет Минздрава России по телемедицине. Затем в марте 2013 года по Указу №325-У была создана система сертификации телемедицинских центров. С 2015 г. в Москве началась реализация Единой информационной системы анализа истории болезней, к концу 2018 года планируется подключить не менее 95% всех государственных медучреждений в стране. Целый ряд проблем, которые мешали развитию этого плана решаются благодаря новому Федеральному закону "Об основах охраны здоровья граждан в Российской Федерации", который вступил в силу 1 января 2018 года. Данный закон предусматривает возможность выписывать рецепты и оказывать медицинскую помощь с использованием телемедицинских технологий путем проведения видеоконсультаций и консилиумов, обеспечивающих взаимодействие врачей и врача и

пациента (или его законного представителя) [1]. На данный момент в России установлено три формата осуществления телемедицинских услуг: врач – пациент, врач – врач (или медицинский работник), местная клиника (поликлиника, медицинский пункт) – центральные органы здравоохранения.

На рынке появляется множество компаний, которые предоставляют телемедицинские услуги. Сейчас в России данными услугами занимаются в основном IT-стартапы, привлекая офлайн-клиники к сотрудничеству. Например, компания «Яндекс» запустила сервис предоставления платных медицинских консультаций, связывающий медиков, пациентов и страховщиков. Девиз сервиса: «Консультируйтесь с врачами, а не на форумах!» Также компания «Мобильные Медицинские Технологии» создает программное обеспечение для онлайн-консультаций «Педиатр 24/7» и «Доктор онлайн». А Сбербанк купил сервис онлайн-записи к врачам docdoc.ru. Пока это только сервис записи к врачам, однако Сбербанк планирует развивать этот сервис дальше и увеличивать количество услуг. Также телемедицину внедряют такие компании, как «Мегафон», «Ростелеком», «МТС» и многие другие [30].

К основным факторам, которые позволяют развиваться телемедицине в России можно отнести [10]:

- **Природно-географические особенности территории**
 - большие территории страны и значительного числа субъектов РФ;
 - Низкая плотность населения на большей части территории страны;
 - Значительное число населенных пунктов, которые находятся на расстоянии от областных центров и не имеют достаточно развитую транспортную инфраструктуру. (В России в сельских населенных пунктах с численностью до 10 тыс. чел. проживают около 13 млн. чел., а с численностью менее 500 человек 9 млн. населения. А всего в сельской местности проживают 37,6 млн человек²);
 - Неблагоприятные условия на территории страны и т.д.
- **Инфраструктурные и организационные особенности системы здравоохранения страны.**
 - Концентрация специализированных медучреждений и высококвалифицированных специалистов в областных, республиканский, краевых центрах;

² Данные Федеральной службы государственной статистики, взятые с сайта gks.ru. Дата обращения: 04.04.2018 г

- Высокие транспортные расходы на поездки в специализированные медучреждения:
- Высокая затратность, потенциальная опасность и организационные сложности транспортировки пациентов из отдаленных районов по санавиации и т.д.

Телемедицина все больше проникает в жизни людей: уже привычным стала онлайн-запись к врачам, многие имеют доступ в личном кабинете к данным с результатами анализов и рекомендациями по лечению. Перспективы впечатляющие, однако существуют и сдерживающие факторы развития телемедицины, а именно недостаток квалифицированных кадров; менталитет населения, не склонный доверять технологиям; проблемы технического характера; нерешенность вопроса о частичном покрытии онлайн-услуг с помощью полисов обязательного медицинского страхования проблемы в нормативно-правовой базе.

Одна из проблем в нормативно-правовой базе - не легитимизирована постановка диагноза с помощью телемедицинских сервисов, то есть для того, чтобы начать пользоваться телемедицинскими сервисами для консультаций, необходимо поставить первоначальный диагноз на очном приеме у врача в поликлинике. Однако по данным Росстата [31] в 2006 году 3,1 процент обратившихся людей старше 15 лет не получают медицинскую помощь при обращении к врачу. Причем были определены следующие причины неполучения надлежащей помощи (Рисунок 1):



Рисунок 1 - Причины отказа от получения медицинской помощи

Как можно заметить из данной диаграммы, 30% опрошенных, которые обращались в поликлинику, не смогли получить первичный диагноз у специалиста в связи с большими очередями и отсутствием записи на прием, а как следствие могут запустить свою болезнь,

что приведет к большим потерям денежных средств, чем при предупреждении болезни. При этом сейчас производится оптимизация врачей, коек и больниц. По данным Росстата [31], численность медицинских работников сократилась на 61 тысячу человек с 2014 по 2015 год (36 тыс. врачей и 25 тыс. человек среднего медицинского персонала). А в 2016 году увеличилось лишь на 7,9 тысяч человек, при этом количество медицинского персонала уменьшилось ещё на 12,2 тысячи человек. Такая тенденция делает медицину все более недоступной и сводится к тому, что медицина в России идет к тому, чтобы лечить запущенные или хронических болезни, а не поддерживать и укреплять здоровье граждан. Поэтому необходимо внедрять технологии, которые помогут предупреждать развитие болезней и контролировать текущее состояние здоровья.

Кроме того, 33% населения в 2016 году не обратились за помощью в медицинские организации вовсе при наличии такой потребности, ссылаясь на следующие причины (Рисунок 2):



Рисунок 2 - Причины не обращения за медицинской помощью при наличии такой необходимости

То есть 32,7 % человек, а это несколько миллионов людей не получили надлежащую медицинскую помощь только из-за того, что у них не было времени добраться до больницы, чтобы поставить первичный диагноз и получить план лечения. Также примерно 49,6% человек не обращаются в поликлиники, потому что не доверяют им свое здоровье и это без учета тех людей, которые лечились самостоятельно по каким-то причинам. За 2016 год количество обращений в клинику было на 47 миллионов меньше. В 2016 году туда ходили 1,204 млрд раз,

в 2015-м — 1,251 млрд, в 2014-м — 1,291 млрд. Данное уменьшение свидетельствует о низкой доступности медпомощи в клиниках.

За 2016 год было зарегистрировано 237,067 млн болезней среди пациентов, обратившихся в поликлинику, при этом 115,187 млн диагнозов было установлено впервые в жизни. При этом количество болезней в долгосрочной перспективе увеличивается. В связи с тем, что вовремя не была проведена профилактика болезни (Рисунок 3).

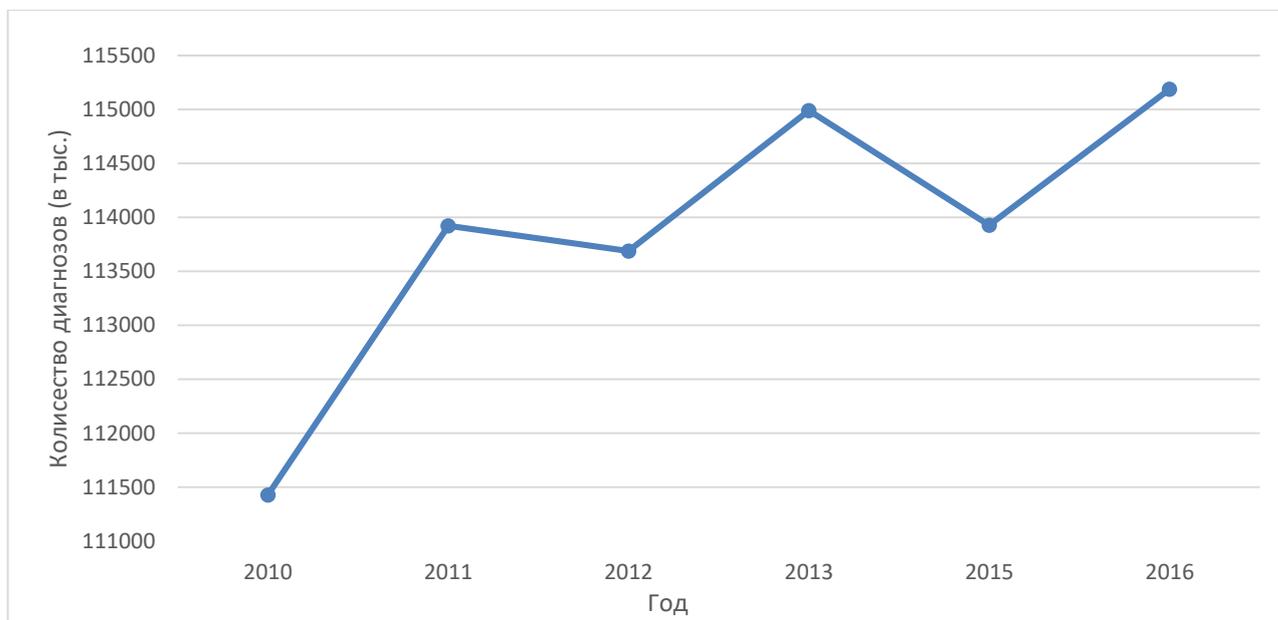


Рисунок 3 - Количество диагнозов, установленных пациентам в первые (в тыс.)

Более того, формально, по данным Фонда обязательного медицинского страхования, 10% медицинской помощи в России оказывается с дефектами, но это всего лишь статистика фонда медицинского страхования. По факту в российских стационарах процент вскрытий умерших пациентов в среднем составляет 60%, причем в 10 - 15% случаев диагнозы, поставленный врачом и выявленный после смерти, расходятся. А если взять тех, кто умер не в стационаре, а дома, то процент расхождения колеблется от 50 до 70%. Неправильно поставленный диагноз сопровождается неправильным лечением³.

Решением данной проблемы может стать создание медицинского приложения на основе нейронных сетей, разработанное при поддержке ведущих врачей России. Данное приложение сможет стать дополнением в развитии телемедицины, так как нейронные сети могут очень точно определять диагнозы и предупреждать возможное появление патологий.

³ Из доклада директора НИИ морфологии человека РАМН Льва Кактурского.

1.3. Использование нейронных сетей в условиях телемедицины

Нейронные сети – это одно из направлений исследований в области искусственного интеллекта, основанное на попытках воспроизвести нервную систему человека. Искусственный нейрон — упрощенная модель биологического нейрона. Данный метод моделирования позволяет воспроизводить сложнейшие нелинейные зависимости. В последние десятилетия это направление становится все популярнее и популярнее (Рисунок 4).

Documents by year

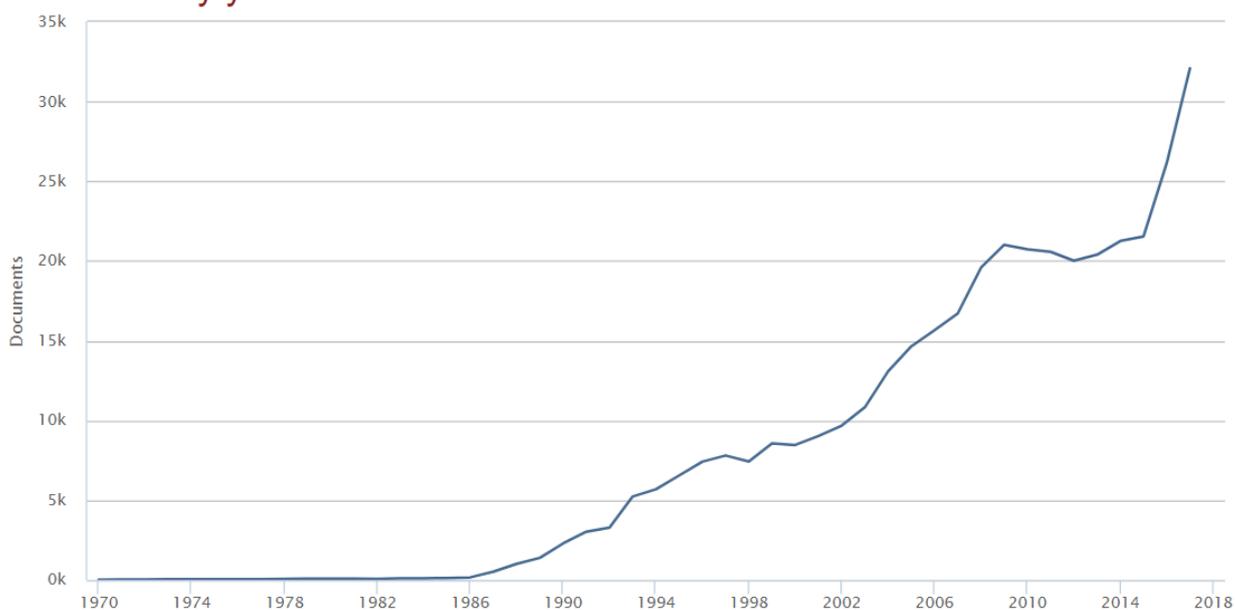


Рисунок 4 - Статистика ежегодных публикаций по теме «нейронные сети» в Scopus

Использование нейронных сетей в медицине началось еще конце 60-х годов XX века и сейчас становится все более и более популярно. Это подтверждается большим числом разработанных медицинских приложений, которые широко применяются в системе здравоохранения многих стран, а также количеством публикаций (только в библиографической и реферативной базе данных Scopus за 2017 год было опубликовано больше 110 тыс. публикаций). При этом количество ежегодных открытий значительно увеличивается (Рисунок 5).

Documents by year

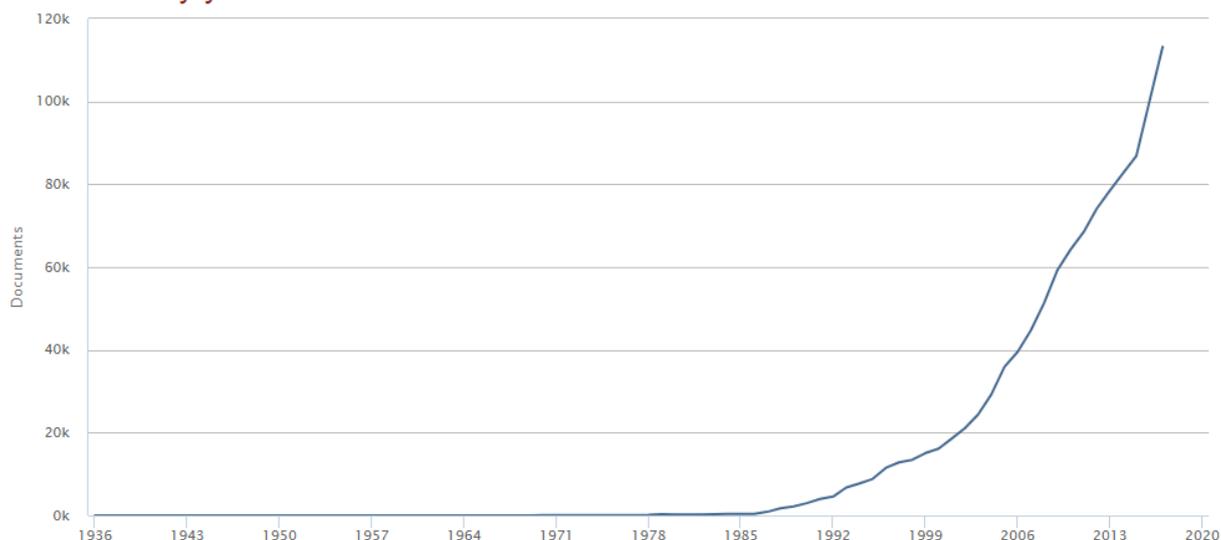


Рисунок 5 - Статистика ежегодных публикаций по теме "Нейронные сети в медицине" в Scopus

Способность к моделированию нелинейных процессов, работе с зашумленными данными и адаптивность дают возможность применять нейронные сети для решения широкого класса различных задач, к ним относятся [3]:

- Классификация образов
Задачей классификации является отнесение каждого наблюдения к одному из нескольких попарно не пересекающихся множеств.
- Кластеризация/категоризация
Задача кластеризации заключается в разбиении множества на группы (кластеры). В каждом кластере должны быть «похожие» по характеристикам объекты, а объекты различных кластеров должны быть отличны друг от друга. Главное отличие кластеризации от классификации заключается в том, что в первом случае перечень групп первоначально не задан и определяется в процессе работы алгоритма.
- Аппроксимация функций
Целью данной задачи является построение аналитической функции, которая наилучшим образом описывает заданную дискретную или непрерывную функцию. С помощью данного построения появляется возможность определить значения функции в других точках, отличных от узловых.
- Предсказание/прогноз
Суть данной задачи состоит в том, чтобы определить значений на несколько периодов вперед значения некоторого наблюдения, зависящего от времени, на

основе предыдущих значений данного наблюдения или других переменных, влияющих на него.

- Оптимизация

Задачей данного алгоритма является поиск такого решения, которое позволит удовлетворить системе ограничений и либо минимизирует, либо максимизирует целевую функцию.

- Память, адресуемая по содержанию

Данная возможность позволяет вызывать содержимое памяти по частичным данным или искаженному содержанию.

Нейронную сеть можно научить выявлять различные патологии лучше, чем способны это делать опытные медсестры и врачи. Машина учится непрерывно 24 часа в сутки, не устает и не болеет. Кроме того, машина ставит диагноз гораздо быстрее, а в большинстве опасных заболеваний время играет решающую роль. Например, если в течении трех часов правильно установить тип инсульта (их более 100), то своевременное лечение позволяет в 90% случаев пациентам вернуться к полноценной жизни. Минуты промедления могут привести к летальному исходу или дорогостоящим процедурам реабилитации. С помощью машинной обработки можно за считанные минуты проанализировать томограмму и правильнее принять решение.⁴

Помимо возможности решать новый ряд задач, нейронные сети обладают и значительным рядом преимуществ, что сделало использование столь популярным. Все преимущества их работы взяты из биологического принципа работы мозга, но, конечно, гораздо упрощеннее по сравнению с ним.

- Распределенная структура.

Скорость обработки информации нейроном ниже на 5-6 порядков, чем производительность кремниевых логических устройств, однако, из-за распределенности и многочисленности взаимосвязей сложно-структурированные и нелинейные задачи мозгом человека выполняются быстрее и эффективнее [3, 11].

- Обучаемость.

Как человек обучается в детстве шаг за шагом, так и нейронная сеть при получении новой информации на обучающем множестве изменяет параметры весов, чтобы добиться минимальной ошибки.

⁴ Тезисы с конференции в «Яндексе», организованной порталом MedAboutMe.
<https://rb.ru/story/it-medicine-in-russia/>

- Адаптивность.

Данная возможность дает нейронным сетям подстраиваться под изменения окружающей среды, перенастраиваться, переобучаться.

- Способность к обобщению.

Нейронная сеть может обрабатывать и получать корректные выходные данные даже если на вход будут поданы данные, выходящие за пределы диапазона обучающей выборки.

- Отказоустойчивость.

Также, как и биологический мозг при повреждении может восстанавливать свою функциональность за счет использования новых взаимосвязей или существующих альтернативных путей, в нейронных сетях также при повреждении сети происходит медленное падение производительности за счет распределенности хранения и обработки информации.

Искусственная нейронная сеть представляет собой совокупность искусственных нейронов и весов связей. Нейрон – это вычислительная единица, которая получает на входе информацию, производит над ней простые вычисления и передает ее дальше. Существуют три основных типа нейронов: входной, скрытый и выходной (Рисунок 6). Совокупность нейронов располагается на слоях. Соответственно, бывают входной, скрытые и выходной слои. Все нейроны оперируют числами только в диапазоне $[0,1]$

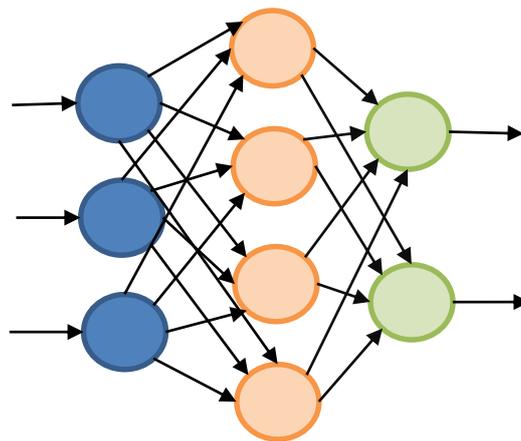


Рисунок 6 - Общий вид нейронных сетей

Каждый нейрон соединяется друг с другом с помощью синапсов. Синапсы обладают одним параметром – весом. Именно благодаря синапсам входная информация изменяется и передается от одного нейрона к другому.

Общую схему работы нейрона можно представить в следующем виде (Рисунок 7). На вход нейрона подаются сигналы x , каждый из которых, проходя через синапсы умножается на вес ω (каждый синапс имеет свой собственный вес), далее производится сложение данных сигналов в сумматоре и добавляется пороговое значение, этот результат преобразуется с помощью функции активации f и является выходом работы нейрона.

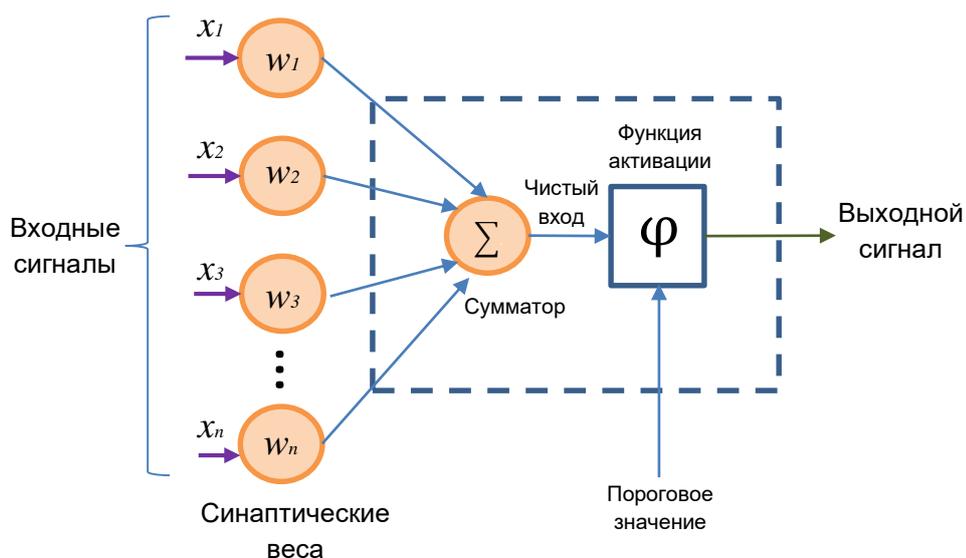


Рисунок 7 - Общая схема работы нейронных сетей

Текущее состояние нейрона описывается формулой:

$$u_i = \sum_{j=1}^N \omega(i, j)x(j) + b(i), \quad (1)$$

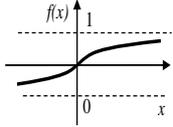
где $x(j), j = 1, 2, \dots, N$ - входные сигналы, величина $b(i)$ называется пороговым значением. $y_i = f(u_i)$ – является выходным сигналом.

Для различных типов искусственных нейронов существуют различные функции активации, которые преобразуют взвешенную сумму в некоторое число, которое и будет является выходом нейрона. Основные функции активации можно увидеть в Таблица 1:

Таблица 1 - Функции активации искусственных нейронов

Название	Формула	Графическая интерпретация
Пороговая	$f(x) = \begin{cases} 1, & \text{при } x \geq \theta \\ 0, & \text{при } x < \theta \end{cases}$	
Знаковая	$f(x) = \begin{cases} -1, & \text{при } x \geq 0 \\ 1, & \text{при } x < 0 \end{cases}$	
Линейная	$f(x) = x$	
Сигмоидальная (логистическая)	$f(x) = \frac{1}{1 + e^{-x}}$	

Продолжение таблицы 1 – Функции активации искусственных нейронов

Гиперболический тангенс	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	
-------------------------	--	---

Чаще всего для работы нейронных сетей используются логистическая функция и гиперболический тангенс, так как они наиболее точно приближены к реальной жизни. Во-первых, данная функция «сжимает» входной сигнал вне зависимости от аргумента до значений от 0 до 1. Во-вторых, значения данных функций принимают не только значения 0 и 1, но и числа, лежащие между ними. Кроме того, во всех точках может быть найдена производная, которая выражается также через эти функции.

В данной работе для построения нейронных сетей будет использоваться пакет для статистического анализа STATISTICA Neural Network, который имеет очень удобный интерфейс и позволяет проводить исследования в диалоговом режиме. Большим преимуществом данного пакета является то, что он встроен в среду аналитических средств программы STATISTICA, что позволяет сочетать как классические, так и нейросетевые методы для получения необходимого результата. Также данный пакет обладает инструментом автоматического нейросетевого поиска, который подбирает наилучшие архитектуры сетей. Такой возможности не предоставляет на данный момент ни один другой пакет. Кроме того, данный пакет включает возможность генерации исходного кода на языках C, C++, C#, Java, PMML, что является немаловажной функцией при создании приложения.

ГЛАВА 2. РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ПЕРВИЧНОЙ САМОДИАГНОСТИКИ ОСНОВАННОГО НА НЕЙРОННЫХ СЕТЯХ

2.1. Описание приложения

2.1.1. Принцип работы приложения

Пациент открывает приложение и в рабочую строку вводит свои симптомы. На основе уже построенной модели система начинает задавать уточняющие вопросы, чтобы получить необходимые данные по всем симптомам. Как только она собрала все данные, пациенту предлагается диагноз, который получается в результате работы нейронной сети. Если после консультации с врачом пациенту был поставлен другой диагноз или выписано другое лекарство, то пациент за некоторые бонусы вносит в систему результат консультации врача, таким образом нейросетевой модуль обучается и впоследствии ставит более точные диагнозы.

Принцип взаимодействия пациента и приложения

- С помощью формы вопрос – ответ предлагает наиболее подходящее лекарство с инструкцией по применению (в качестве вопросов могут предлагаться картинки, анимации);
- Для каждого лекарства указана средняя цена, а также на карте указаны местоположения ближайших аптек, имеющих данное лекарство;
- Функция проверки совместимости действия препаратов;
- Рекомендация плана действий и врачей-специалистов, к которым необходимо обратиться, с указанием их стоимости и местонахождения;
- Наличие услуги онлайн-консультации врачей;
- В личном кабинете есть возможность хранить результаты всех анализов, исследований, используемых лекарств и результатов консультаций с врачами, которые могут использоваться врачом и пациентом для ведения истории болезней.

Вопросы пациенту организовываются в формате простых вопросов, фотографий, анимаций, видеороликов. Ответы пациентов реализуются в формате «Да/Нет», в формате вопросов по шкале боли от 0 до 10 и открытые численные ответы (Рисунок 8).



Рисунок 8 - Прототип интерфейса приложения: виды вопросов и рекомендации.

2.1.2. Построение информационно-логической модели (ИЛМ)

Результирующей таблицей работы приложения является таблица – Поставленный диагноз и рекомендации. Пациенту выводится область, к которой относится болезнь, и специалисты в этой области, к которым необходимо обратиться. Также предоставляется перечень лекарств, которые могут помочь и адреса аптек, где можно купить эти лекарства. Все данные берутся из справочников (Рисунок 9).

Одно лекарство может быть использовано для лечения различных болезней, каждая болезнь может быть побеждена различными лекарствами. Связь Лекарства и Болезни – *многие-ко-многим*. Связь происходит с помощью таблицы Лекарства-Болезни, связующие реквизиты – код лекарства и код болезни.

Одно лекарство может находиться во многих аптеках, в каждой аптеке находится множество лекарств. Связь Лекарства – Аптеки – *многие-ко-многим*. Связь осуществляется с помощью таблицы Наличие лекарства, связующие реквизиты – код лекарства и код аптеки.

Каждая болезнь относится к одной области медицины, к одной области медицины относятся множество различных болезней. Связь – *один-ко-многим*, *связующий реквизит* – Код специализации.

Каждый врач относится к одной области специализации, в одной области может быть множество специалистов. Связь Врачи и Области медицины - *один-ко-многим*. Связующий реквизит – Код специализации.

Онлайн-специалисты, врачи, к которым пациенты могут обратиться за консультацией дистанционно, относятся к одной области медицины, в то время как в одной области может быть множество онлайн-специалистов. Связь Онлайн-специалисты – Области медицины – *один-ко-многим*. Связующий реквизит - Код специализации.

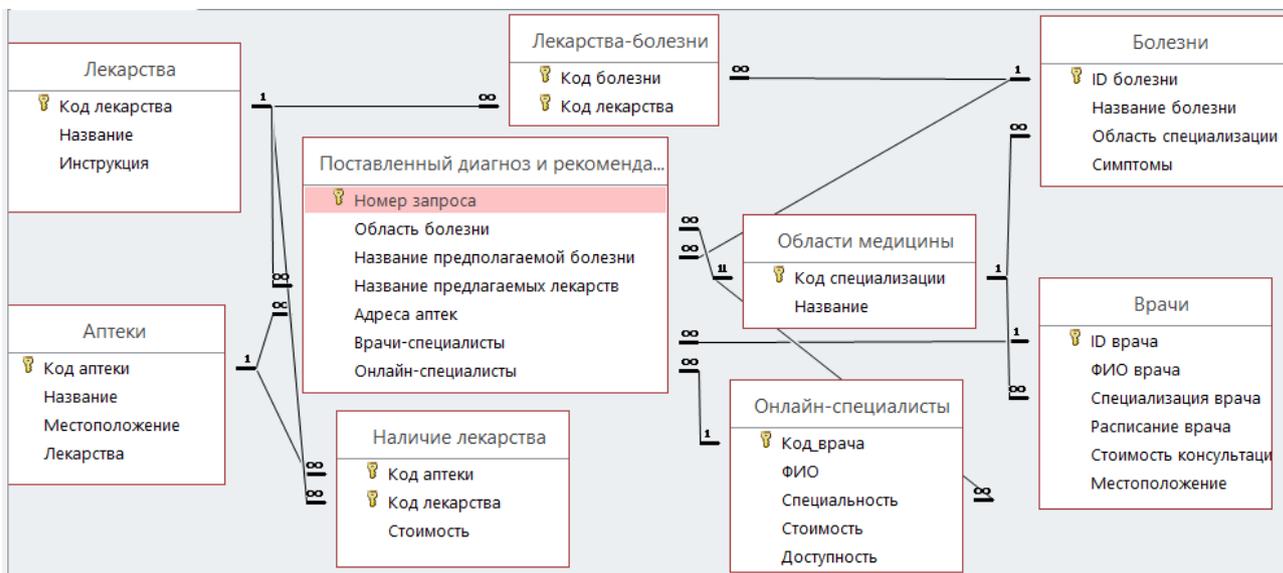


Рисунок 9 - Информационно-логическая модель приложения

2.2. Реализация нейросетевого модуля

Для построения прототипа нейросетевого модуля и анализа его качества были выбраны пять близких по симптоматике болезней, относящихся к одной области, но имеющих различное лечение:

- Катаральный гастрит;
- Аппендицит;
- Пиелонефрит (Воспаление почек);
- Энтерит (Воспаление тонкого кишечника);
- Отравление.

Симптомы катарального гастрита: колющие и режущие боли вокруг пупка и в верхней части живота, вздутия, тяжесть в эпигастрии, отрыжка, изжога, тошнота, неприятный привкус во рту, возможна температура и рвота с примесями желчи, слабость, головокружение, белый налет на языке.

Симптомы аппендицита: острая боль с правой стороны или в районе пупка; твердый, в некоторой степени даже жесткий живот; высокая температура; озноб и сильное

потоотделение; запор и, возможно, рвота или понос, с присутствием частичек крови; тошнота или рвот; боли в спине, ложные позывы к испражнению.

Симптомы пиелонефрита: боль в пояснице, температура, головная боль, частое и болезненное мочеиспускание, слабость.

Симптомы энтерита: частый жидкий стул, тошнота и рвота, боли около пупка, повышенная температура, урчание и вздутия, общее недомогание и головные боли, белый налет на языке, признаки гиповитаминоза, снижение массы тела.

Симптомы отравления: повышение температуры, боль в желудке, вздутие, отрыжка, понос, рвота, слабость, понижение давления, спазмы в животе.

Для того, чтобы можно построить нейросетевой модуль, необходима статистика. Так как согласно статье 13 Федерального закона⁵ «сведения о факте обращения гражданина за оказанием медицинской помощи, состоянии его здоровья и диагнозе, иные сведения, полученные при его медицинском обследовании и лечении, составляют врачебную тайну», нет возможности получить медицинские карты с симптомами и поставленными диагнозами для построения модели. Поэтому с целью получения достоверных данных были изучены врачебные пособия [22, 23, 25] и была сгенерирована случайным образом статистика с помощью наложенных на симптомы ограничений. (Таблица 2)

Таблица 2 - Ограничения на симптомы для выбранных болезней

Болезнь	Т °С	Живот с п.с.	Живот с л.с.	Живот центр	Спина п.с	Спина л.с	Стул	Рвота	Аппетит	Боль при моч..
Аппендицит	37,5-39	6-10	0-4	4-10	3-8	0-5	Диарея Запор	Да Нет	Нет	0-1
Гастрит	36,6-38	4-10	1-4	3-8	0-2	-> 0	Диарея Запор Норм.	Да Нет	Да Нет	0
Отравление	36,6-39	3-9	0-4	0-5	0-2	-> 0	Диарея	-> нет	-> нет	0
Пиелонефрит	36,6-39,5	0-4	0-4	0-6	1-10	1-10	Норм.	Да Нет	Да Нет	1-9
Этерит	37-39,5	2-6	2-6	3-10	0-3	0-3	Диарея	Да	-> нет	0-1

⁵ Федеральный закон от 21.11.2011 N 323-ФЗ "Об основах охраны здоровья граждан в Российской Федерации"

Полученные результаты были загружены в пакет для статистического анализа Statistica. Также была добавлена дополнительная переменная кодов для обучающей и контрольной выборок. На каждую болезнь было предоставлено 30 обучающих наблюдений и 9 тестовых.

Таблица 3 - Загруженные данные в пакет для статистического анализа Statistica

1 Type	2 Illness	3 Temperature	4 Pain_right _front	5 Pain_left _front	6 Pain_middle _front	7 Pain_right_ back	8 Pain_left_ back	9 appetite	10 stool	11 vomiting	12 pain_urina tion
Train	Appendicitis	38,2	7	4	4	8	1	0	1	1	0
Train	Appendicitis	38,0	9	1	5	8	0	0	1	0	0
Train	Appendicitis	38,4	10	4	8	5	1	0	1	0	0
Select	Appendicitis	38,6	9	0	7	8	5	0	2	1	0
Select	Appendicitis	37,9	7	4	9	3	2	0	1	0	0
Select	Appendicitis	38,7	9	2	10	5	1	0	1	0	1
Select	Appendicitis	38,7	9	3	8	4	1	0	2	0	1
Select	Appendicitis	37,6	10	1	7	3	5	0	1	0	1
Select	Appendicitis	38,8	9	2	8	7	4	0	1	0	0
Select	Appendicitis	38,4	8	3	10	5	2	0	2	1	1
Select	Appendicitis	38,7	8	3	7	8	0	0	2	1	0
Select	Appendicitis	38,6	7	2	6	3	4	0	2	1	0
Train	enterit	39,3	3	6	8	0	1	0	1	1	1
Train	enterit	38,5	3	3	8	3	1	0	1	1	1

Стадии решения задач с помощью нейронных сетей

В задачи классификации сеть должна отнести каждое наблюдение к одному из нескольких классов. В пакете Statistica Neural Networks американской фирмы Statsoft для классификации используется номинальная выходная переменная – различные ее значения соответствуют различным классам. Данный пакет относится к современным универсальным нейросетевым продуктам и поэтому более совершенен по сравнению с ранее выпущенными. Он имеет мощные алгоритмы обучения сети (включая методы сопряженных градиентов и Левенберга-Маркара), возможность создания сложных комбинаций из сетей различных архитектур. Для этого пакета характерны простота в использовании и аналитические мощности, например, Automatic Network Designer (автоматический конструктор сети) определит наилучшую архитектуру для конкретной задачи, осуществит отбор переменных

1) Выбор типа анализа

Данная задача относится к задаче классификации, пакет Statistica позволяет автоматически выбирать ее (Рисунок 10).

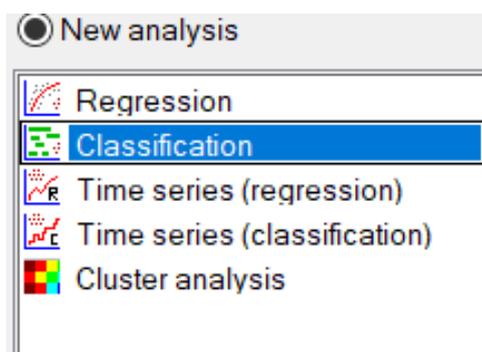


Рисунок 10 - Типы анализа в пакете

2) Выбор стратегии построения нейронных сетей

В Statistica предлагается 3 варианта построения нейронных сетей:

- Автоматизированная нейронная сеть (АНС);
- Пользовательская нейронная сеть (ПНС);
- Метод многократных подвыборок.

Для построения своей модели была выбрана АНС, потому что в этом случае программа будет в автоматическом режиме перебирать *различные* типы и архитектуры сетей и выбирать из них наилучшие модели. При выборе ПНС программа строит сети только *одной* архитектуры, которые полностью задаются пользователем. В последнем случае будет построен ансамбль сетей *одной* архитектуры, который строится по разным подвыборкам.

Далее была выбрана категориальная целевая переменная – болезни, и непрерывные входные переменные – все симптомы болезней (Рисунок 11).

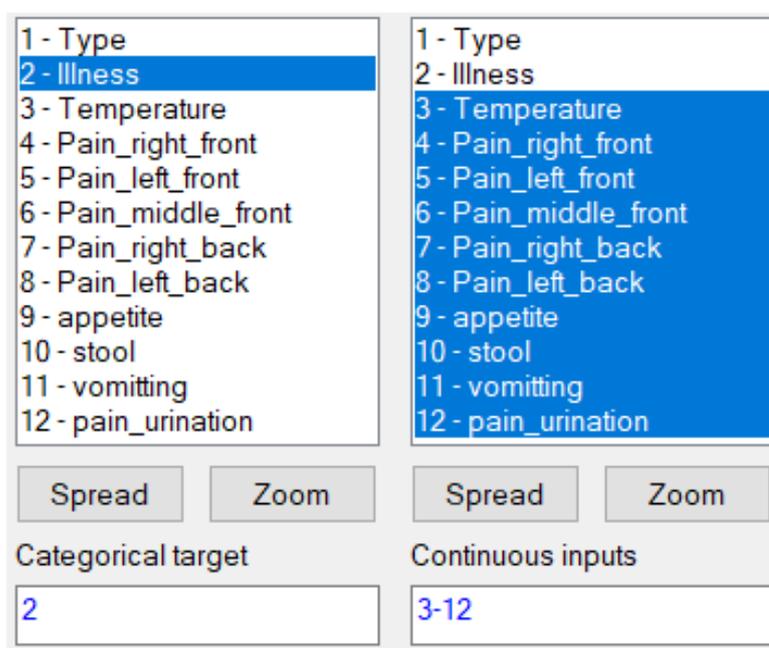


Рисунок 11 - Выбор категориальной и непрерывных входных переменных

Затем были заданы коды для подвыборок, где обучающее множество - данные со значением Train в переменной Type, а контрольное – данные со значением Select в переменной Type (Рисунок 12).

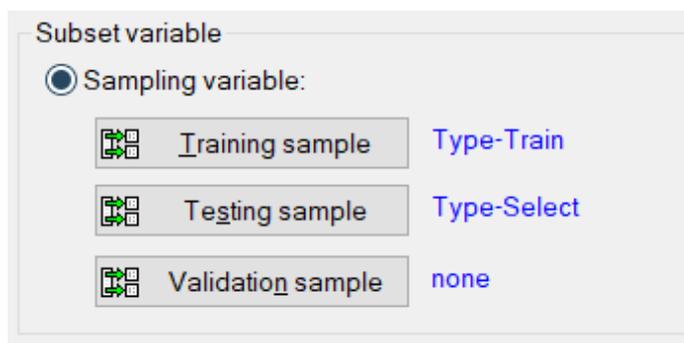


Рисунок 12 - Выбор кодов для подвыборки

3) Задание параметров перед обучением сетей

На этом этапе были выбраны функции активации, так как это модель классификации я выбрала логическую и гиперболическую как для скрытых нейронов, так и для выходных.

Далее необходимо определить количество скрытых слоев и скрытых нейронов на этих слоях. Выбор правильного количества является очень важным шагом, потому что если будет слишком мало нейронов, то сеть может не обучиться хорошо. Но если нейронов слишком много, то сеть может переобучиться, то есть на обучающей выборке она будет предоставлять отличные результаты, однако плохо на входных примерах, не входящих в обучающее множество.

Существует эвристический способ нахождения скрытых слоев и нейронов на них. Функция данной задачи не определена на конечном множестве точек, поэтому одного скрытого слоя будет мало, но большинство функций могут быть аппроксимированы с помощью четырехслойного перцептрона (хотя существуют сложные реальные задачи, которым требуется намного больше слоев). Количество нейронов на данных слоях вычисляются по следующим формулам:

$$r = \sqrt[3]{\frac{m}{n}} \quad (2)$$

$$k_1 = mr^2 \quad (3)$$

$$k_2 = mr, \quad (4)$$

где m – число нейронов на входном слое, n – число нейронов на выходном слое, k_1 – число нейронов в первом скрытом слое, k_2 – число нейронов на втором скрытом слое.

В данной модели 10 входных нейронов (симптомы) и 5 выходных нейронов. Таким образом количество скрытых нейронов определяется:

$$r = \sqrt[3]{\frac{10}{5}} \approx 1,26 \quad (5)$$

$$k_1 = 10 * (1,26)^2 \approx 16 \quad (6)$$

$$k_2 = 10 * 1,26 \approx 13 \quad (7)$$

Однако в настоящее время нет никаких жестких правил, каким образом выбирается точное количество скрытых слоев и нейронов. Невозможно определить хорошую топологию сети основываясь только на количестве входов и выходов, поэтому модель была обучена на различных минимальном и максимальном количестве скрытых нейронов (Рисунок 13).

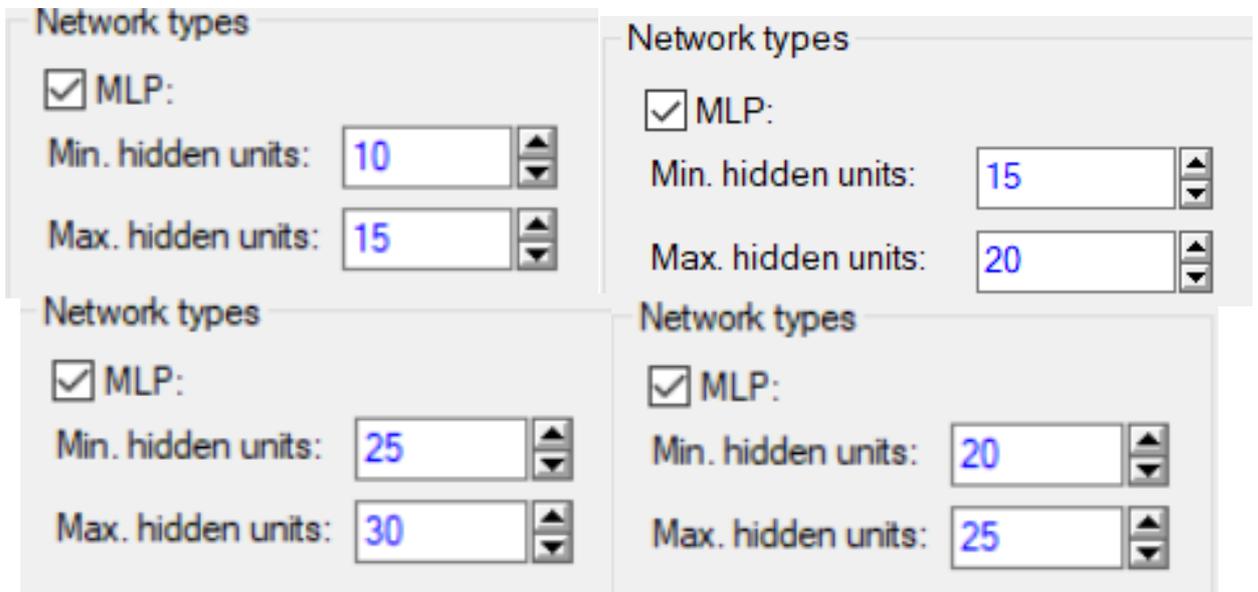
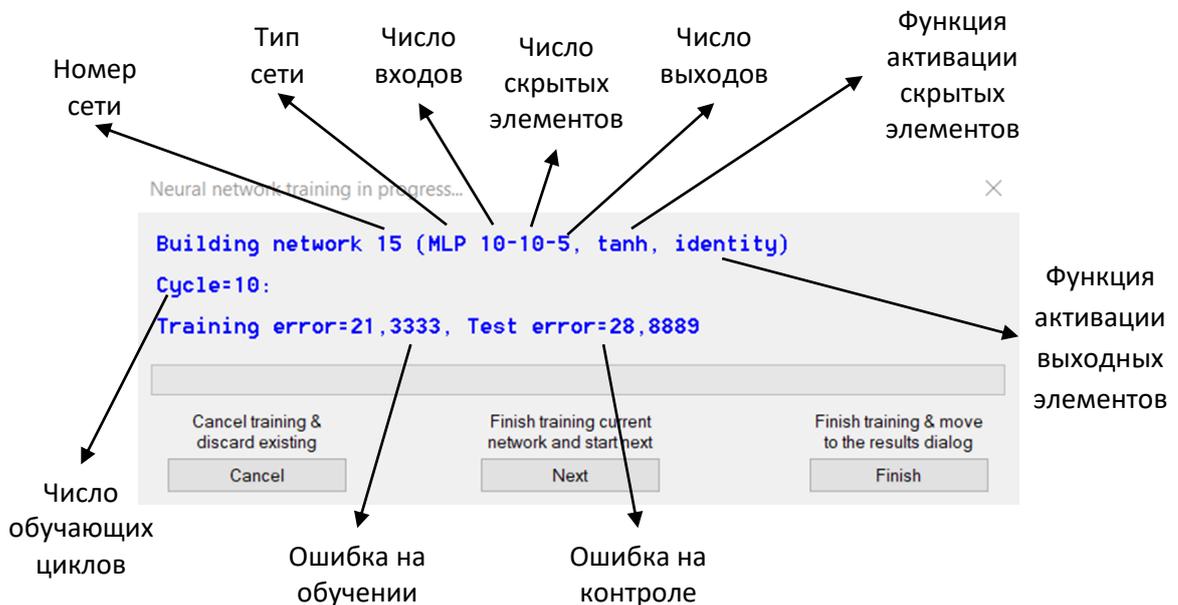


Рисунок 13 - Выбор минимального и максимального числа скрытых нейронов

4) Обучение сетей



5) Анализ результатов

Наилучшие результаты достигаются в случае 23 и 24 скрытых элементов, функции ошибки – кросс-энтропия и сумма квадратов, входная функция активации – логистическая в обоих случаях, выходные – многомерная логистическая и экспоненциальная.

Таблица 4 - Анализ полученных моделей

Summary of active networks (Spreadsheet1)								
Index	Net. name	Training perf.	Test perf.	Validation perf.	Training algorithm	Error function	Hidden activation	Output activation
1	MLP 10-12-5	96.9697	100.0000		BFGS 25	Entropy	Logistic	Softmax
2	MLP 10-14-5	98.1818	100.0000		BFGS 49	SOS	Tanh	Tanh
3	MLP 10-15-5	97.5758	100.0000		BFGS 63	SOS	Exponential	Identity
4	MLP 10-12-5	94.5455	100.0000		BFGS 33	SOS	Exponential	Identity
5	MLP 10-11-5	93.9394	100.0000		BFGS 21	SOS	Logistic	Logistic
6	MLP 10-18-5	97.5758	100.0000		BFGS 53	SOS	Logistic	Identity
7	MLP 10-20-5	96.3636	100.0000		BFGS 22	SOS	Tanh	Logistic
8	MLP 10-16-5	96.9697	100.0000		BFGS 56	SOS	Exponential	Exponential
9	MLP 10-17-5	98.7879	100.0000		BFGS 20	Entropy	Tanh	Softmax
10	MLP 10-20-5	98.1818	100.0000		BFGS 48	SOS	Logistic	Exponential
11	MLP 10-23-5	100.0000	100.0000		BFGS 33	Entropy	Tanh	Softmax
12	MLP 10-24-5	100.0000	100.0000		BFGS 26	SOS	Tanh	Logistic
13	MLP 10-21-5	98.1818	100.0000		BFGS 55	SOS	Exponential	Identity
14	MLP 10-23-5	98.1818	100.0000		BFGS 52	SOS	Tanh	Identity
15	MLP 10-23-5	98.1818	100.0000		BFGS 45	SOS	Tanh	Exponential
16	MLP 10-26-5	96.3636	100.0000		BFGS 46	SOS	Tanh	Identity
17	MLP 10-27-5	95.1515	100.0000		BFGS 35	SOS	Tanh	Exponential
18	MLP 10-30-5	96.9697	100.0000		BFGS 17	Entropy	Logistic	Softmax
19	MLP 10-30-5	97.5758	100.0000		BFGS 50	SOS	Tanh	Tanh
20	MLP 10-30-5	92.7273	100.0000		BFGS 11	Entropy	Identity	Softmax

Однако из-за большого количества слоев возможно переобучение нейронных сетей, что сказывается на том, что сеть будет очень хорошо работать на обучающем и контрольном множествах, однако на тестовом выдаст множество ошибок. Поэтому для анализа были выбраны 3 модели с наилучшими результатами.

Таблица 5 - Матрица сопряженности

		Illness-Appendicitis	Illness-Gastrit	Illness-enterit	Illness-poisoning	Illness-pyelonephritis	Illness-All
9.MLP 10-17-5	Total	33.0000	33.0000	33.0000	33.0000	33.0000	165.0000
	Correct	33.0000	33.0000	33.0000	31.0000	33.0000	163.0000
	Incorrect	0.0000	0.0000	0.0000	2.0000	0.0000	2.0000
	Correct (%)	100.0000	100.0000	100.0000	93.9394	100.0000	98.7879
	Incorrect (%)	0.0000	0.0000	0.0000	6.0606	0.0000	1.2121
11.MLP 10-23-5	Total	33.0000	33.0000	33.0000	33.0000	33.0000	165.0000
	Correct	33.0000	33.0000	33.0000	33.0000	33.0000	165.0000
	Incorrect	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Correct (%)	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000
	Incorrect (%)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
12.MLP 10-24-5	Total	33.0000	33.0000	33.0000	33.0000	33.0000	165.0000
	Correct	33.0000	33.0000	33.0000	33.0000	33.0000	165.0000
	Incorrect	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Correct (%)	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000
	Incorrect (%)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Модель 9 ошиблась дважды в отнесении отравления к нужному классу.

Анализ ошибки. Принцип отнесения к классу основывается на том, что победитель берет все. Так как модель 9 считает, что эти симптомы относятся к гастриту с вероятностью 44,7%, а к отравлению с вероятностью 55,3%, модель на выходе дает гастрит.

Таблица 6 - Анализ полученной ошибки

Illness - Output 9.MLP 10-17-5	Illness-Appendicitis 9.MLP 10-17-5	Illness-Gastrit 9.MLP 10-17-5	Illness-enterit 9.MLP 10-17-5	Illness-poisoning 9.MLP 10-17-5	Illness-pyelonephritis 9.MLP 10-17-5
poisoning	0.000000	0.000043	0.000000	0.999957	0.000000
Gastrit	0.000000	0.552990	0.000084	0.446926	0.000000

2.3. Анализ качества модели

Первоначально был проведен анализ чувствительности. Данный анализ позволяет увидеть, какие параметры наиболее и наименее важны для каждой модели. Например, для модели 12 самым важным показателем является боль при мочеиспускании, а все остальные не так важны. Разница в важности значительна, это означает, что при дальнейшем тестировании мы могли бы убрать все значения, кроме боли при мочеиспускании. Такая модель плохая, скорее всего она переобучена, поэтому она была убрана из дальнейшего анализа.

Таблица 7 - Анализ чувствительности моделей

Networks	pain_urination	vomitting	stool	Pain_right_back	Pain_middle_front	Pain_right_front	Pain_left_back	Pain_left_front	Temperature	appetite
9.MLP 10-10-5	16,666	161,5653	24,7970	117,2680	89,42262	77,64396	55,08031	48,30029	64,08390	19,98604
11.MLP 10-20-5	55,913	53,0596	08,0026	40,3374	27,47720	26,18972	43,62506	39,92811	13,59829	7,68756
12.MLP 10-19-5	6317,609	1,4911	71,4496	1,6967	1,90041	5,10183	0,92345	2,54961	1,27240	0,67219
Average	2130,063	72,0386	68,0830	53,1007	39,60008	36,31184	33,20961	30,25934	26,31820	9,44860

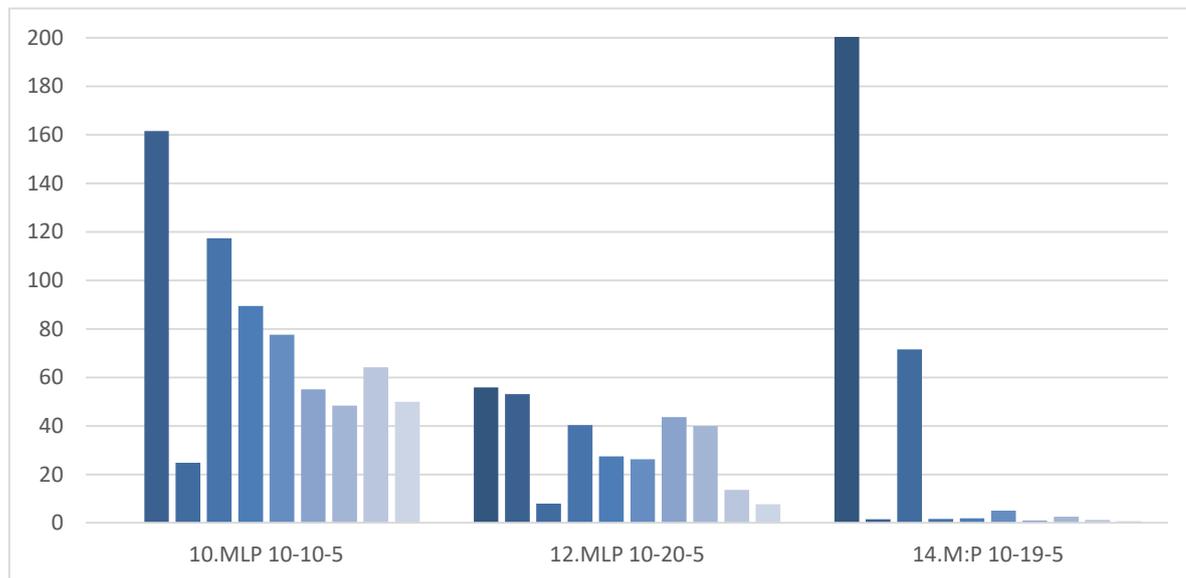


Рисунок 14 - Диаграмма анализа чувствительности модели

Затем был проведен анализ полноты оставшихся моделей, которая показывает отношение верно угаданных объектов класса ко всем представителям этого класса. Для модели 11 все показатели равны 1. Для модели 9 каждой болезни, кроме отравления этот показатель равен 1, для отравления данный показатель равняется 0,93 (Рисунок 15).

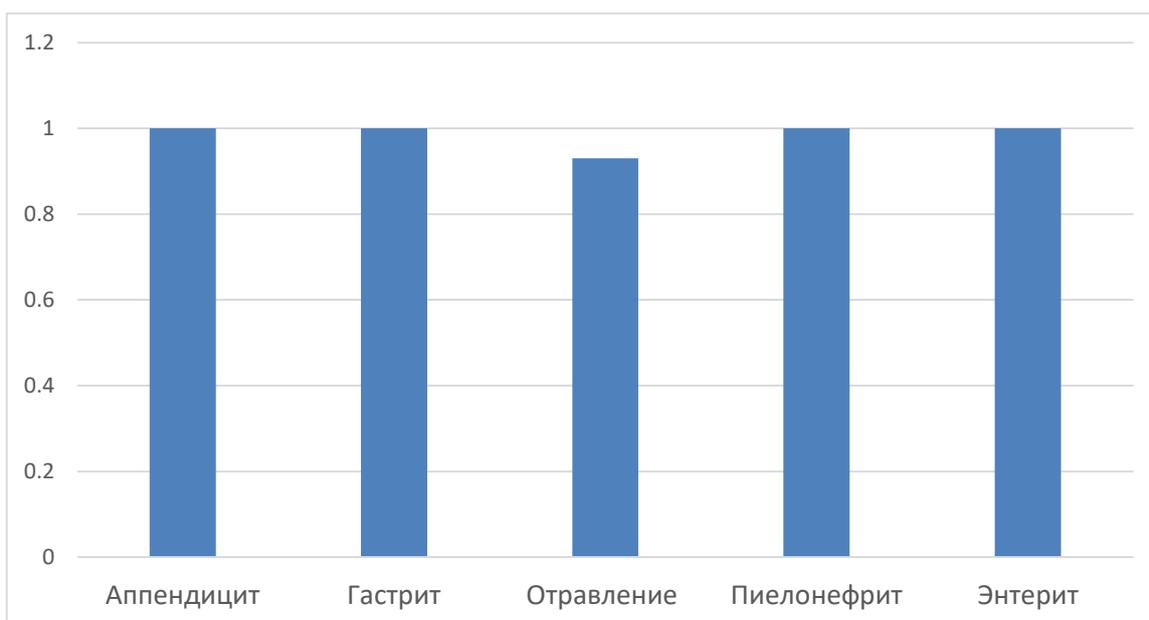


Рисунок 15 - Анализ полноты модели 9

Также был проведен анализ точности (Таблица 7), который показывает отношение верно угаданных объектов класса ко всем объектам, которые мы определили, как объекты класса. Можно заметить, что данный результат оказался достаточно высоким.

Таблица 8 - Анализ точности модели 9

Аппендицит	1
Гастрит	0,971
Отравление	1
Пиелонефрит	1
Энтерит	0,972

Последним этапом тестирования была проверка правильности определения болезни на новых значениях. Для этого были добавлены по пять новых наблюдения и было проведено сравнение входных данных с выходными результатами двух моделей (Таблица 9). В модели 9 правильный результат выводился в 96% случаев. Для такой небольшой статистики результат ошибки в 4% достаточно хороший. В модели 11 лишь в 76% случаев выводится правильный результат, что подтверждает тот факт, что модель была все же переобучена, поэтому давала для обучающего множества правильность ответов 100%, а уже на тестовой проверке 76%.

Таблица 9 - Результаты тестирования модели

Target_illness	10.Illness_(t)	12.Illness_(t)
Gastrit	Gastrit	poisoning
Gastrit	Gastrit	Gastrit
Gastrit	Gastrit	enterit
Gastrit	Gastrit	Gastrit
Gastrit	Gastrit	Gastrit
Appendicitis	Appendicitis	enterit
Appendicitis	Appendicitis	Appendicitis
Appendicitis	Appendicitis	pyelonephritis
Appendicitis	Appendicitis	Appendicitis
Appendicitis	Appendicitis	enterit
poisoning	poisoning	pyelonephritis
poisoning	poisoning	poisoning
poisoning	Gastrit	poisoning
poisoning	poisoning	poisoning
poisoning	poisoning	poisoning
pyelonephritis	pyelonephritis	pyelonephritis
enterit	enterit	enterit

Для наглядности вида полученной нейронной сети был построен граф нейронной сети в Deductor.

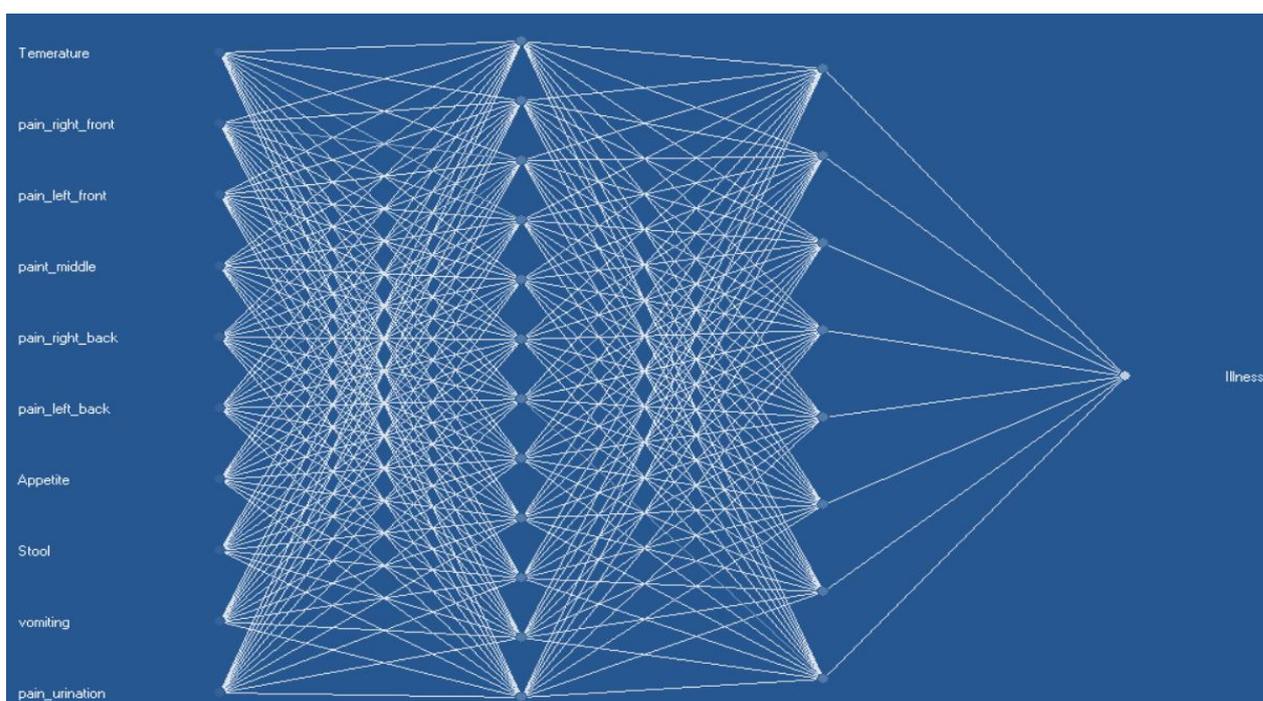


Рисунок 16 - Граф нейронной сети

2.4. Оценка стоимости разработки программного обеспечения

Оценка стоимости разработки программного обеспечения оценивалась по модели конструктивных затрат COSOMO (Constructive Cost Model) базового уровня. Данная модель наиболее широко используется для оценивания, потому что она основана на регрессии и оценивает фактический размер, понесенные трудозатраты, а также фактическую длительность разработки ПО. Данная модель впервые была разработана Бари Боэмом в 1981 году, как результат анализа проектов компании «TRW Aerospace», она основывалась на водопадной модели жизненного цикла разработки ПО, однако позднее в 1997 году была усовершенствована и стала более подходящей для оценки современных проектов разработки ПО, получив название COSOMO II.

Преимущества данной модели заключаются в том, что она достаточно простая, потому что включает в себя небольшое количество данных для определения трудозатрат и стоимости. Следовательно, это статически однозначная модель, потому что формулы были получены из тщательного анализа большого количества проектов, таким образом, можно утверждать, что она близка к реальности. Факторы, которые используются для оценки, очень полезны в понимании формирования стоимости проекта. Кроме того, данная модель легко интерпретируема за счет простых методов оценки.

Однако данная модель имеет также и ряд недостатков. Во-первых, это эмпирическая модель оценки, а, следовательно, она создана на основе ограниченного количества проектов. Это увеличивает вероятность получения неточных результатов, так как различные модели зависят от неточных констант. Во-вторых, данная модель не учитывает навыки разработчиков, знания, взаимодействия, вопросы аппаратного обеспечения, уровень текучести персонала и другие параметры. Также данная модель упрощает влияние аспектов настройки безопасности.

Данная модель применяется к трем классам проектов разработки ПО [14]:

- Органический вид (Organic mode) – относительно небольшие и простые программные проекты (обычно не больше 50 тыс. строк), обладающие небольшой проектной группой, члены которой имеют опыт разработки подобных проектов, и нежестким набором требований.
- Полуинтегрированный вид (Semi-detached mode) – разрабатываемые проекты обычно сложнее, чем в органическом виде. Некоторые члены команды могут

иметь опыт работы с такими продуктами, остальные могут быть и не знакомы с аспектами разрабатываемой системы. Программный продукт содержит не больше 300 тыс строк. (Например, Системы обработки транзакций с фиксированными требованиями к оборудованию терминала и обслуживанием баз данных)

- Встроенный вид (Embedded mode) – существуют жесткий набор требований на аппаратное, программное и операционное обеспечение. (Например, программное обеспечение управления полетом авиационной техники)

Базовая модель COSOMO предоставляет приблизительную оценку параметров проекта, на данном уровне рассчитывается трудоемкость и стоимость разработки как функция от размера программы. Размер выражается в оценочных тысячах строк.

Трудоемкость вычисляется с помощью формулы:

$$E_{\text{(ожидаемая)}} = a \cdot (SIZE)^b \text{ (человеко – месяцев)}, \text{ где} \quad (8)$$

SIZE - размер продукта (количество строчек кода в 1000), коэффициенты *A* и *E* – константы, зависящие от класса продукта (Таблица 10).

Срок разработки (длительность) находится по формуле:

$$D = c \cdot E^d \text{ (месяцев)}, \text{ где} \quad (9)$$

E – трудоемкость, показатели *c* и *d* – константы, определяющиеся классом проекта (Таблица 10).

Число разработчиков определяется на основе полученных результатов:

$$N = \frac{E}{D} \text{ (человек)} \quad (10)$$

Таблица 10 - Значения коэффициентов в зависимости от класса проекта

Класс проекта	a	b	c	D
Органический	2.4	1.05	2.5	0.38
Полуинтегрированный	3.0	1.12	2.5	0.35
Встроенный	3.6	1.2	2.5	0.32

Разработка приложения, основанного на нейросетях, относится к полуинтегрированному классу проектов, потому что она имеет множество как жестких требований, так и нет (Например, к жестким требованиям относится защита клиентской базы,

которая охраняется законодательством), кроме этого, она включает многие аспекты, которые ранее еще не были в России реализованы.

Для определения количества строк кода были взяты данные о количестве строк у приложения Babylon Health (приложение разработанное британской стартап компанией Alí Parsa), которое имеет относительно похожие характеристики и функции и имеет 1.1 млн строк кода. Основываясь на этом факте, была посчитана трудоемкость и срок реализации:

$$E_{\text{(ожидаемая)}} = 3 \cdot (1100)^{1.12} \approx 7647 \text{ (человеко – месяцев)} \quad (11)$$

$$D = 2.5 \cdot 7647^{0.35} \approx 57 \text{ (месяцев)} \quad (12)$$

Средняя заработная плата разработчика программного обеспечения на 2015 год по данным Росстата равнялась 44894 рубля. Темп роста реальных заработных плат был 0,8% в 2016 году относительно 2015 и 3,5% в 2017 относительно 2016. Таким образом, с поправкой на этот темп роста на начало 2018 года средняя заработная плата программиста составляет примерно 46837 рублей. Таким образом разработка данного приложения будет стоить примерно:

$$P = 7647 \cdot 46837 = 357,881 \text{ млн рублей} \quad (13)$$

2.5. Результаты внедрения приложения

Если сравнивать показатели качества модели, то при использовании приложения, основанного на нейронных сетях, показатель точности результатов значительно выше, как в сравнении с данными Фонда обязательного медицинского страхования, так и с более пессимистичными данными статистики вскрытия.

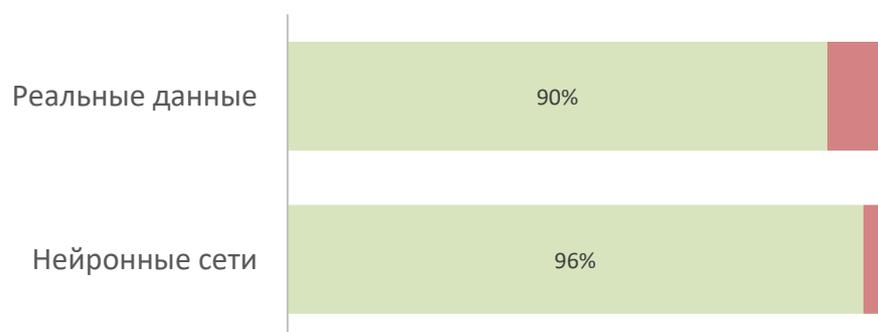


Рисунок 17 - Сравнение показателей качества двух моделей

Значительное количество населения не уделяет достаточного внимания предупреждению болезней. Пациенты обычно откладывают профилактику в сторону и

пытаются сосредотачиваться на актуальной проблеме. Это связано с тем, что сбои могут происходить в различных системах нашего организма и поэтому пациентам проще фокусироваться на решении появившихся проблем, чем предотвратить их появление. Однако около 73% представителей здравоохранения и 72% граждан (Рисунок 17), опрошенных в 19 странах мира, включая Россию, считают, что цифровые технологии очень важны для проведения профилактики [27].

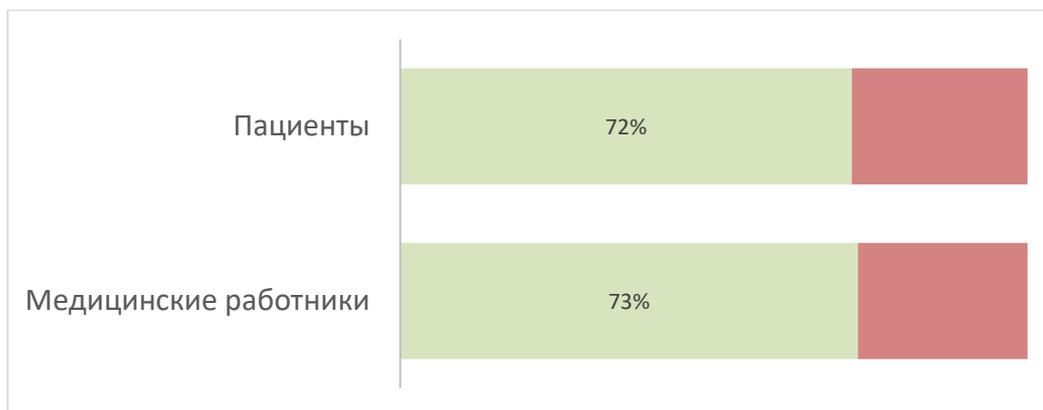


Рисунок 18 - Процент людей, считающих, что цифровые технологии необходимы для профилактики

Кроме того, чтобы создать эффективную систему здравоохранения и обеспечить ее устойчивое развитие, необходимо сместить фокус с лечения на проведение профилактики. Стоимость первичного обращения по полису ОМС к врачу-терапевту составляла на 2017 год 108,22 рубля⁶. Однако если вовремя болезнь не остановить и запустить ее таким образом, что будет необходима госпитализация, то стоимость уже будет составлять 861,03 рубля, если медицинская помощь оказывается в условиях дневного стационара, и 8781,82 рубль, если медицинская помощь оказывается в стационарных условиях медицинскими организациями. Если же рассматривать проведение операций, то суммы будут еще более значительные

Ранее было посчитано, что стоимость разработки приложения составляет 357,881 млн рублей за 4,5 года разработки. Таким образом, чтобы окупилась его разработка, необходимо, чтобы им воспользовались хотя бы 3,306 млн раз:

$$\frac{357\,881\,000}{108.22} = 3,306 \quad (14)$$

В первой главе рассматривалась статистика обращений пациентов в поликлиника и было выявлено, что многие пациенты не идут в поликлинику к врачу, пока у них что-либо не произойдет, потому что у них нет времени добираться до осмотра, ни доверия врачам. Однако большинство из них готовы пользоваться инструментами и технологиями на базе

⁶ Тарифы на оплату медицинской помощи, оказываемой в амбулаторных условиях в рамках Территориальной программы ОМС, применяемые, в том числе для осуществления горизонтальных расчетов // Приложение № 6 к Тарифному соглашению на 2017 год от «29» декабря 2016 года

искусственного интеллекта, а также сервисами дистанционных консультаций. Поэтому данная технология достаточно быстро окупится, а также принесет значительный вклад в развитие здравоохранения в России. Кроме того, если эти граждане смогут вовремя установить диагноз, у государства будет возможность сохранить деньги на устранение последствий запущенной болезни.

В Америке использование телемедицинских средств для наблюдения и предупреждения болезней снизило количество попаданий в стационары на 19%, а продолжительность нахождения в поликлиниках сократилась на 25% [20]. Средняя продолжительность пребывания пациента в стационаре на койке за 2016 год составляло 11,1 дней, при сокращении на 25% за счет использования средства предупреждения данная величина может сократиться до 8,4 дней. То есть в среднем выгода будет в 23 710 рублей с человека. А если количество попаданий в больницу сократиться на 19%, то выгода будет 288 725 тысяч рублей на 100 человек населения.

Таким образом, разработка и внедрение данного приложения не только улучшит качество жизни населения страны, но также сможет сократить количество расходов на лечение граждан и перераспределить их на улучшение предоставляемых услуг и на профилактику заболеваний.

ЗАКЛЮЧЕНИЕ

В данной дипломной работе рассматривалась возможность и перспективы разработки приложения для постановки первичного самодиагноза пациентов, основанного на нейронных сетях. Для этого была проанализирована сфера здравоохранения в России и были выявлены проблемы для дальнейшего ее усовершенствования. Далее была разработана инфологическая модель приложения и создан его прототип в пакете для статистического анализа STATISTICA Neural Network, который предоставил высокие показатели полноты и точности отнесения симптомов к правильной болезни. Заключительным этапом данной работы было сравнение существующей модели здравоохранения в России с созданной во время исследования моделью.

Оценка затрат на разработку приложения осуществлялась с помощью модели конструктивных затрат СОСОМО II. Стоимость разработки оценивается в размере 357,881 млн рублей, а длительность разработки в 57 месяцев. Данное приложение обладает быстрыми сроками окупаемости, потому что стоимость одного приема терапевта на данный момент составляет 108,22 рубля, следовательно, приложением необходимо воспользоваться как минимум 3,306 млн человек, а это меньше 1% количества всех обращений к врачам, при этом в этом проценте не учитываются люди, которые не смогли или не захотели прийти на осмотр по причинам отсутствия времени или другим причинам. Кроме того, использование данного приложения может сократить количество попаданий в стационар, за счет предупреждения болезней, а также продолжительность пребывания в больнице.

СПИСОК ЛИТЕРАТУРЫ

1. Федеральный закон от 21.11.2011 (ред. 15.06.2017) N 323-ФЗ "Об основах охраны здоровья граждан в Российской Федерации"
2. Федеральный закон от 21.11.2011 N 323-ФЗ "Об основах охраны здоровья граждан в Российской Федерации"
3. Абрагин А.В. ПЕРСПЕКТИВЫ РАЗВИТИЯ И ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ [Книга]. - Иваново : Олимп, 2015.
4. Алиев Х.Р. Модель планирования и управления разработкой сложных программных систем на основе комбинированной методики оценки трудозатрат [Отчет]. - Санкт-Петербург : СПбГУ, 2010.
5. Баранов А.А, Вишнева Е.А., Намазова-Баранова Л.С. Телемедицина - перспективы и трудности перед новым этапом развития [Статья] // Педиатрическая фармакология. - [б.м.] : ПедиатрЪ, 2013 г.. - 3 : Т. 10.
6. Боровиков В.П. Нейронный сети. Statistica Neural Networks. Методология и технологии современного анализа данных. [Книга]. - Москва : Горячая линия - Телеком, 2008.
7. Владзимирский А.В. ОЦЕНКА ЭФФЕКТИВНОСТИ ТЕЛЕМЕДИЦИНЫ [Книга]. - Донецк : Цифровая типография, 2007.
8. Исаев Т.М. МЕТОДЫ ОЦЕНКИ ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ ТЕЛЕМЕДИЦИНЫ [Статья] // Экономика и управление. Вопросы экономики и права. - Москва : [б.н.], 2012 г..
9. Леванов В.М. ОТ ТЕЛЕМЕДИЦИНЫ ДО ЭЛЕКТРОННОГО ЗДРАВООХРАНЕНИЯ: ЭВОЛЮЦИЯ ТЕРМИНОВ [Статья] // МЕДИЦИНСКИЙ АЛЬМАНАХ. - Нижний Новгород : Ремедиум Приволжье, 21.04.2012 г.. - 2.
10. Эдириппулиге С., Ожегова Л.А., Ожегов А.Ю. ФАКТОРЫ РАЗВИТИЯ И СОВРЕМЕННОЕ СОСТОЯНИЕ ТЕЛЕМЕДИЦИНЫ: ГЕОГРАФИЧЕСКИЙ АСПЕКТ [Статья] // География. Геология. - 2017 г.. - 3 : Т. 3.
11. Ярушкина Н.Г Основы теории нечетких и гибридных систем [Книга]. - Москва : Финансы и статистика , 2004.

12. Adnan ATAÇ Engin KURT, S.Eray YURDAKUL 13th International Educational Technology Conference [Конференция] // An Overview to Ethical Problems in Telemedicine. - New York : Elsevier Ltd, 2013.
13. Al-Shayea Qeethara Kadhim, Artificial neural network in medical diagnosis [Статья] // International journal of computer science. - 2011 г..
14. Boehm Barry, Software cost estimation with COCOMO II [Книга]. - New Jersey : Prentice Hall , 2000.
15. Burns Elizabeth A., Kenneth Korn, James Whyte IV, Oxford American Handbook of CLINICAL EXAMINATION AND PRACTICAL SKILLS [Книга]. - [б.м.] : Oxford University Press Inc, 2011.
16. Coiera Enrico, Guide to Medical Informatics, the Internet, and Telemedicine [Статья] // Books, Journals, New Media. - New York : [б.н.], 1997 г..
17. Dalia S. Ashour Dina M. Abou Rayia, Mohamed Maher Ata, Amira S. Ashour, Mustafa M. Abd Elnaby, Dr. Peter Saggau, Hybrid feature extraction techniques for microscopic hepatic fibrosis classification [Статья] // Wiley Periodicals. - 2018 г..
18. Danica Mitch M. Pacis Edwin D. C. Subido, Nilo T. Bugtai, AIP Conference Proceedings [Конференция] // Trends in telemedicine utilizing artificial intelligence. - [б.м.] : American Institute of Physics, 2018.
19. Eberhart Russell C. , Dobbins Roy W., Neural network PC tools: a practical guide [Книга]. - [б.м.] : Academic Press Inc., 1990.
20. Eysenbach Gunther, Deborah Greenwood, Azizeh Sowan, and Elizabeth Krupinski, Personalized Telehealth in the Future: A Global Research Agenda [Статья] // Journal of Medical Internet Research. - [б.м.]: JMIR Publications, 2016 г.
21. Harno K. Telemedicine in managing demand for secondary care services [Статья] // Telemed and Telecare. - 1999 г..
22. Houghton Andrew R, Gray David, SYMPTOMS AND SIGNS IN CLINICAL MEDICINE. An Introduction to Medical Diagnosis. [Книга]. - London : Edward Arnold (Publishers) Ltd, 2010.
23. Mohan Harsh, Textbook of Pathology [Книга]. - New Delhi : Jaypee Brothers Medical Publishers (P) Ltd, 2010.
24. OmniconHealthGroup, The Power of Artificial Intelligence [Журнал]. - [б.м.] : OmniconHealthGroup, 2016 г..

25. Papadakis Maxine A., Stephen J. McPhee, CURRENT Medical Diagnosis & Treatment [Книга]. - San Francisco : McGraw Hill Education, 2017.
26. Pearce G., L. Mirtskhulava, K. Bakuria, J. Wong, S. Al-Majeed, N. Gulua 17th UKSIM-AMSS International Conference on Modelling and Simulation [Конференция] // Artificial Neural Network and Mobile Applications in Medical diagnosis . - 2015.
27. Philips, Future health index 2017. Care that delivers. - [б.м.] : Philips, 2017 г..
28. Varrecchia T., De Marchis, C., Rinaldi, M., Draicchio, F., Serrao, M., Schmid, M., Conforto, S., Ranavolo, A. Lifting activity assessment using surface electromyographic features and neural networks [Статья] // International Journal of Industrial Ergonomics. - July 2018 г.. - стр. 1-9.
29. World Health Organization TELEMEDICINE // Opportunities and developments in Member States. - Geneva : [б.н.], 2010 г..
30. ТЕЛЕМЕДИЦИНА.RU Телемедицина в России [В Интернете] // ТЕЛЕМЕДИЦИНА.RU первое профильное СМИ. - 2018 г.. - URL: <https://telemedicina.ru/> (дата обращения: 15.02.2018 г.).
31. Федеральная служба государственной статистики ЗДРАВООХРАНЕНИЕ В РОССИИ [Отчет]. - Москва : Росстат, 2017. URL: http://www.gks.ru/free_doc/doc_2017/zdrav17.pdf (дата обращение 17.03.2018 г.).
32. Laing Drew, Telemedicine in healthcare benefits physicians and patients [В Интернете] // Nextech.com. - Drew Laing, 24 06 2015 г. - URL: <http://www.nextech.com/blog/telemedicine-possess-bright-future-in-healthcare> (дата обращения 03.04.2018 г.).
33. URL: <https://proxy.library.spbu.ru:2092> (дата обращения 02.12.2017 г.)