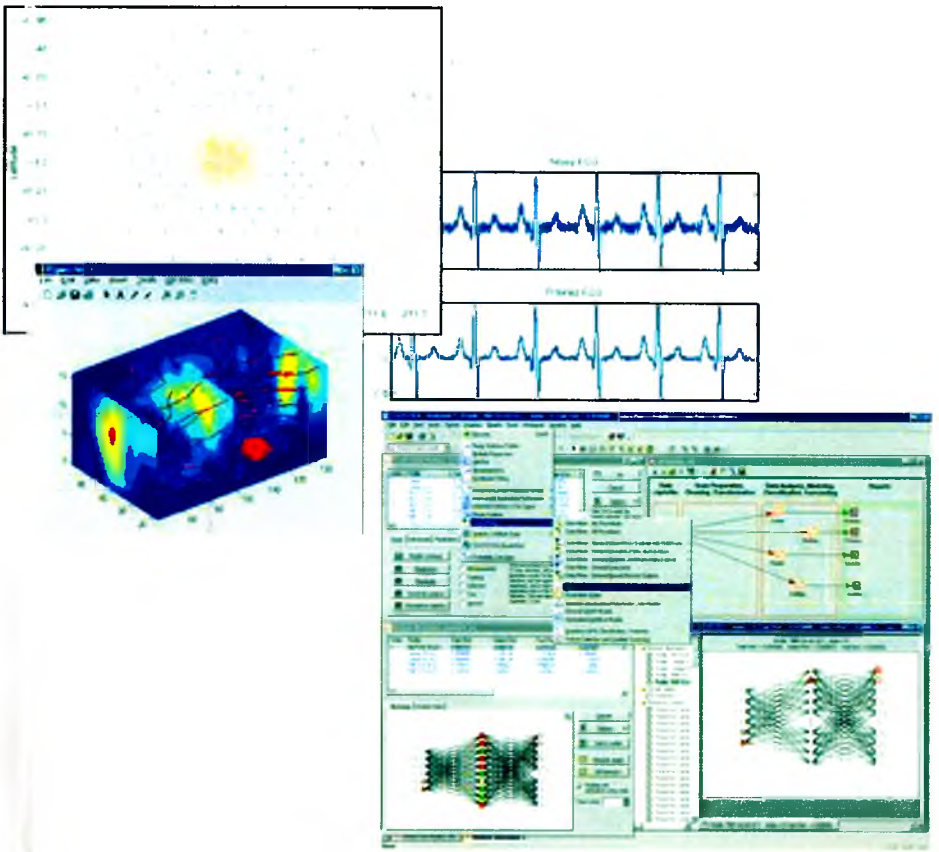
Т.М. МАГРУПОВ, С.А. ВАСИЛЬЕВА, М.Т. МАГРУПОВА

АНАЛИЗ И ОБРАБОТКА МЕДИКО-БИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ



Министерство высшего и среднего специального образования
Республики Узбекистан

Ташкентский государственный технический университет
им. Абу Райхана Беруни

Магрупов Т.М., Васильева С.А., Магрупова М.Т.

АНАЛИЗ И ОБРАБОТКА МЕДИКО-БИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

Ташкент -2012

УДК 519.23:578.087.1
ББК Р 34с11 + В 172.6

Магрупов Т.М., Васильева С.А., Магрупова М.Т. Анализ и обработка медико-биологической информации. –Т.: ТашГТУ, 2012. –152 с.

Монография посвящена актуальным вопросам в области биомедицинской инженерии с применением основных положений современных информационных технологий.

Описаны задачи медико-биологических исследований с использованием методов статистического анализа; приведены результаты применения нейросетевых методов для анализа и обработки медико-биологических данных, основанные на использование пакета STATISTICA Neural Networks; исследованы проблемы визуализации и обработки медицинских сигналов и изображений с учетом широких возможностей системы MATLAB. Приведены результаты исследований применения информационных технологий в медицине. На основе разработанных методов и алгоритмов предложена экспертная система для решения задач дифференциальной диагностики, прогнозирования, выбора стратегии и тактики лечения в отоларингологии. Предложены способы анализа и обработки данных процессов метаболизма в организме и функций внешнего дыхания, а также методика автоматизированной системы ведения истории болезни.

Теоретические положения проведённых исследований сопровождаются примерами и иллюстративными материалами.

Монография предназначена научным работникам и специалистам в области биомедицинской инженерии, а также медикам и биологам, занимающимся вопросами применения информационных технологий в медицине и биологии.

Утверждено к печати Учёным советом Ташкентского государственного технического университета им. Абу Райхана Беруни Министерства высшего и среднего специального образования РУз.

Отв. редактор: д.т.н., проф. Ш.М. Гулямов

Рецензенты: д.м.н., проф. Х.Э. Шайхова, Ташкентская медицинская академия;
д.т.н., проф. Х.З. Игамбердиев, Ташкентский государственный технический университет

© Ташкентский государственный технический университет, 2012
©© Магрупов Т.М., Васильева С.А., Магрупова М.Т., 2012

Введение

Характерной особенностью математизации и компьютеризации медицины и биологии в наши дни является стремительный рост спроса на такие методы обработки эмпирического материала, которые обеспечивают комплексный подход к познанию живых организмов. В таких исследованиях всегда учитывается принцип единства и взаимосвязанности явлений в природе.

Одним из обязательных этапов любого биомедицинского исследования является статистический анализ данных. Продолжительное время анализ медицинских данных был уделом специалистов, так как это требовало серьезной предварительной подготовки. С появлением и совершенствованием современных программ обработки данных статистическая обработка поднялась на новый уровень. Теперь исследователь-медик может и не иметь математической подготовки. Достаточно оперировать статистическими понятиями и, самое главное, правильно выбрать метод анализа. Все осуществимо благодаря компьютеру и новейшим программам.

Использование статистических методов анализа данных позволяет выявлять объективные взаимосвязи показателей биологических объектов, в частности живого организма, в условиях присутствия случайных факторов и делать на их основе обоснованные выводы и прогнозы.

Такие математические методы, которые разработаны с всесторонним учетом принципа единства живой природы и возможности практической их реализации и использования программного обеспечения, являются достижением в области постановки и анализа медико-биологических исследований. Однако опыт показывает, что и в век информационной технологии лучших успехов достигают специалисты, умеющие не только использовать обработанную информацию, но также уяснить сущность применяемых методов. Это предохранит от механического их

использования, которое рано или поздно приводит к нелепым или даже абсурдным выводам.

К сожалению, сегодня уровень представления результатов статистического анализа данных медико-биологических исследований все еще остается недостаточно высоким.

Ясно, что многообразие задач статистического анализа данных медико-биологических исследований, обусловленное разнообразием целей и методов исследований, не позволяет создать универсальный алгоритм подобного анализа, одинаково пригодный во всех практических ситуациях. Однако к настоящему времени накоплен обширный опыт проведения статистического анализа данных и выработаны практические рекомендации, выполнение которых является обязательным элементом любого статистического выборочного исследования.

Статистические методы заняли прочные позиции в арсенале современной медицины и фармакологии. Практически нет такого метода статистического анализа, который не нашел бы применения в медицине. В клиниках, например, они служат важным вспомогательным средством для получения информации о влиянии различных факторов на распространенность и течение заболеваний, а также дают возможность сравнивать эффективность различных методов диагностики, лечения и т.д. Кроме того, бывают случаи, когда только путем анализа статистических данных можно определить, являются ли некоторые побочные эффекты следствием применения конкретного препарата или неправильной постановки диагноза.

Последовательность типовых действий, выполняемых в процессе статистического анализа результатов, успешно реализовывается на базе статистических пакетов, в частности, пакета прикладных программ (ППП) STATISICA фирмы StatSoft Inc. (США). Выбор ППП STATISICA для иллюстрации методов анализа авторами монографии не случаен, поскольку он обладает достаточными возможностями для решения основных задач медицинских исследований.

Задачи медицинской диагностики и прогнозирования не имеют четких (явных) алгоритмов решения. В условия таких задач входит большое число сложно комбинирующихся факторов. Способ решения этих задач человеком лишь в малой степени основан на четких правилах. В основном используется опыт (явная или неявная память о предыдущих ситуациях), подразумевающий правильное решение не только в случае повторения ситуации, но и при возникновении совершенно новой, не встречавшейся ранее ситуации.

Повысить эффективность медицинской диагностики и прогнозирования может позволить математический аппарат искусственных нейронных сетей (ИНС), которые способны обобщать эмпирические данные, выявлять и воспроизводить скрытые в этих данных закономерности, что дает возможность рассматривать нейросетевой подход как наиболее приемлемую альтернативу классическим статистическим методам.

ИНС, работая по неявным алгоритмам и решая задачи, не имеющие явного решения, достаточно хорошо моделируют способ принятия решений человеком. ИНС, являясь перспективной технологией обработки и обобщения больших объемов медицинской информации для решения задач классификации и прогнозирования, представляют собой нелинейную систему, позволяющую классифицировать данные гораздо лучше, чем обычно используемые линейные методы. В приложении к медицинской диагностике ИНС являются одним из основных инструментов поддержки принятия решений врачом-специалистом в условиях отсутствия точных моделей реальных процессов и явлений.

Доступность и возросшие вычислительные возможности современных компьютеров привели к широкому распространению программ, использующих принципы нейросетевой обработки данных.

В качестве успешного применения современных компьютерных технологий в медико-биологических исследованиях можно назвать использование таких систем, как MATLAB, использующих различные

высокоэффективные методы обработки биомедицинских сигналов и изображений. Информационные сигналы, имеющие место в различных медико-биологических исследованиях, весьма разнообразны. Одно из центральных мест среди них занимают изображения внутренних органов человека, правильная интерпретация которых в случае патологических процессов приводит к принятию обоснованного решения врачом на всех этапах проведения лечебно-диагностических процессов.

Построение экспертных систем (ЭС) относится к ещё одной важной области исследований, посвященной формализации способов представления знаний. Разработка медицинских экспертных систем для решения медицинских задач является особенно актуальной, поскольку в большинстве случаев они представлены большим объемом многомерных, запутанных, а порой и противоречивых клинических данных.

В медицинском программном обеспечении отдельную группу представляют программы для интенсивной терапии, которая использует оценочные системы, позволяющие проводить сравнение пациентов, оказавшихся в критических состояниях, и оценивать результаты их лечения.

В эту же группу входят расчетно-диагностические программы для использования в работе терапевта и врача-анестезиолога, которые должны содержать современные расчетные данные; наиболее полно, объективно и быстро выдавать результаты; не создавать трудностей при их использовании.

В современных условиях лечение одного заболевания может представлять сложный комплекс из множества госпитализаций (в том числе в разных медицинских учреждениях) и амбулаторного наблюдения, при этом заводится множество историй болезни и амбулаторных карт, доступ к которым для лечащего врача весьма затруднен даже внутри одной организации. Становится актуальным развитие электронной истории болезни (электронной карты) пациента. В основу принципа работы составления, учета и хранения медицинской

информации в электронном виде, заложена идея создания единого информационного ресурса, который позволяет оперировать с личными данными пациентов, а также обмениваться такими данным с другими медицинскими учреждениями.

К настоящему времени существует несколько подходов к решению проблемы разработки и внедрения в лечебно–профилактические учреждения (ЛПУ) электронной истории болезни. Применение компьютерных технологий позволяет создавать электронную модель такого объекта, как «Медицинская карта стационарного (амбулаторного) больного» в интересах различных пользователей и в разных целях. В идеале, такая модель должна устраивать всех заинтересованных лиц и обеспечить повышение качества всех процессов управления здоровьем пациентов.

Представленная авторами работа посвящена исследованиям и разработке вышеперечисленных положений информационных технологий в различных областях медицины для теоретического и практического анализа и обработки медико-биологических данных.

Авторы выражают глубокую признательность зам. министру здравоохранения РУз д.м.н., проф. Хожибекову Марату Худайкуловичу и генеральному директору Республиканского специализированного центра хирургии им. В.Вахидова д.м.н., проф. Назирову Ферузу Гафуровичу за содействие в организации проведения научных исследований и внедрение результатов в медицинские учреждения, а также д.м.н., проф. Шайховой Халиде Эркиновне и к.м.н. Закирову Кадыру Насыровичу за многолетнее творческое сотрудничество и всестороннюю поддержку в ходе подготовки монографии.

ГЛАВА 1. МНОГОФАКТОРНЫЙ КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ ДАННЫХ МЕДИЦИНСКИХ ИССЛЕДОВАНИЙ

В настоящее время становится всё более актуальным корректное применение статистических методов, научный подход к планированию медицинских исследований. Основная цель применения статистических методов – сведение к минимуму случайных ошибок в научном медицинском и клинических исследованиях. Прежде всего в медицине статистика используется в задачах, связанных с выборочными обследованиями, с проверкой эффективности различных доз лекарственных средств, диагностикой заболеваний на основании проводимых медицинских анализов, прогнозированием выздоровления больных и т.д. Возможности для решения указанных задач предоставляет ППП STATISTICA фирмы StatSoft Inc. (США). Широкий спектр статистических методов модулей ППП STATISTICA так или иначе используется в медицине.

1.1. Задачи исследований сложных систем

Известно, что объекты исследования в медицине представляют собой сложные вероятностные (стохастические) системы. Сложные системы функционируют при воздействии на них множества входных факторов. Часть из них является контролируемыми X_1, X_2,..., X_k, измеряемыми количественно или оцениваемыми в баллах. Другая часть входных факторов относится к группе неконтролируемых, случайных факторов: они не поддаются измерению, но оказывают воздействие на систему, результатом которого является случайность ее функционирования. Состояние системы характеризуется множеством выходных параметров Y_1, Y_2, ..., Y_l, которые также измеряются количественно или в баллах и представляют собой случайные величины, следующие нормальному или

иному закону распределения с соответствующими числовыми характеристиками. Наилучшие результаты многомерного статистического анализа данных медицинских исследований получают тогда, когда распределение входных факторов и выходных параметров нормальное или близкое к нему.

Наблюдавшиеся значения k факторов и l параметров для n объектов сводятся в матрицу наблюдений размером $n\times(k+l)$. По матрице наблюдений с помощью ППП STATISTICA 6.0 проводятся:

- статистическое описание переменных;
- корреляционный анализ;
- канонический корреляционный анализ;
- регрессионный анализ.

В результате статистического описания устанавливают закон распределения переменных и определяют их числовые характеристики, строят графики основных зависимостей между факторами и параметрами.

Корреляционный анализ обеспечивает оценку связей всех переменных попарно.

Канонический корреляционный анализ даёт оценку связи всего множества входных факторов со всеми выходными параметрами в совокупности.

На основе канонического корреляционного анализа можно судить о достаточности связи входных факторов, включенных в матрицу наблюдений, и выходных параметров, характеризующих состояние системы.

Моделирование каждого выходного параметра методами регрессионного анализа даёт возможность построить линейные или нелинейные модели, которые используются для решения основных задач системного анализа:

- изучения характера изменения выходных параметров при изменении входных факторов;
- оценки степени влияния факторов на параметры;

- прогнозирования параметров при заданных значениях факторов;

- поиска оптимальных уровней факторов для получения требуемых значений параметров;

- оценки информативности параметров при заданной совокупности воздействующих факторов.

Требования к базе данных для многомерного статистического анализа. Матрица наблюдений с n строками по числу наблюдавшихся объектов в выборке и $(k+l)$ столбцами по числу наблюдавшихся k входных факторов и l выходных параметров должна содержать только количественные данные в натуральных единицах измерения или баллах.

При отсутствии данных по какому-либо признаку его заменяют средним значением признака для всей выборки, хотя это приводит к искажению исходной информации. Следует также иметь в виду, что некоторые статистические пакеты не рассчитывают корреляционной матрицы в случае, когда число переменных превышает число наблюдений. Надежное решение можно получить, если в матрицах наблюдений число строк n в 3-5 раз превышает число столбцов $(k+l)$.

Все данные должны быть тщательно проверены: устраняются грубые ошибки, исключаются явно анормальные результаты наблюдения. Выборка должна быть, безусловно, репрезентативной по отношению к исследуемой генеральной совокупности.

В соответствии с целью и задачами исследования в матрицу необходимо ввести дополнительные столбцы с группированными признаками, например, группированный признак G1 – контрольная группа с кодом 1, опытная группа с кодом 2; группированный признак пола G – мужчины с кодом 1, женщины с кодом 2 и т.п.

Задачи и содержание многомерного корреляционного анализа. Многомерный корреляционный анализ проводится для количественной оценки направления, силы и значимости линейной связи между всеми переменными базы данных попарно. Такая связь характеризуется коэффициентом корреляции Пирсона.

В результате решения по опциям Descriptive statistics и Correlation на экран выводятся следующие результаты:

- таблица числовых характеристик переменных;

- корреляционная матрица, содержащая коэффициенты корреляции и уровни их значимости для всех пар переменных.

По таблице числовых характеристик анализируется соответствие распределений каждой переменной нормальному закону.

По корреляционной матрице, представляющей собой квадратную симметричную таблицу с размером $(k+l)\times(k+l)$, судят о направлении, силе и значимости корреляционной связи переменных попарно, в особенности о связи входных факторов с выходными параметрами.

Вариант расчета числовых характеристик переменных и корреляционной матрицы, а также интерпретации результатов приведены в подглаве 1.2.

Назначение и содержание канонического корреляционного анализа. Канонический корреляционный анализ предназначен для изучения связи между входными факторами и выходными параметрами в их совокупности.

Для проведения канонического корреляционного анализа в исходной матрице наблюдений с размерами $n\times(k+l)$, где n – число наблюдавшихся объектов, k – число входных факторов и l – число выходных параметров, выделяют две группы переменных:

1. Left set – группа выходных параметров;

2. Right set – группа входных факторов.

Алгоритмом предусмотрено:

1. Определение ограниченного числа канонических переменных обобщающих выходные параметры l-ой группы, и такого же количества канонических переменных, обобщающих входные факторы 2-ой группы. При этом первая пара канонических переменных обобщает наибольшую часть дисперсии переменных, вторая пара - большую долю из оставшейся части дисперсии и т.д. Количество пар канонических переменных зависит

от размерности матрицы наблюдений. Практика показала, что 2-3 пар канонических переменных достаточно для надежного представления всей совокупности переменных.

2. Формирование полей рассеяния объектов в координатах первой, второй, третьей пары канонических переменных для 1-ой и 2-ой группы, а после их формирования - расчет канонических коэффициентов корреляции: Can r_1 - по паре первых, Can r_2 - по паре вторых, Can r_3 - по паре третьих канонических переменных.

По величине канонических коэффициентов корреляции судят о силе связи между совокупностями входных факторов и выходных параметров. Квадраты коэффициентов (Eigen value) характеризуют степень детерминации совокупности параметров совокупностью факторов для каждой пары канонических переменных. Значимость канонических коэффициентов корреляции детерминации оценивают по χ^2-критерию Пирсона. Коэффициенты считают значимыми при вероятности равной и более 0,95 или при уровне значимости $p<0,05$;

3. Расчёт факторной структуры канонических переменных (Factor structure), т.е. коэффициентов корреляции, характеризующих направление и силу корреляционной связи канонических переменных с наблюдавшимися входными факторами и выходными параметрами. В результате дается оценка важности входных факторов и информативности выходных параметров.

Такой анализ на начальном этапе исследования позволяет оценить достаточность связи между входными факторами и выходными параметрами с целью построения для них надежных моделей, а также выделить наиболее значимые факторы и информативные параметры откликов на воздействия.

Назначение и содержание многомерного регрессионного анализа. Построение линейного уравнения регрессии. Многомерный регрессионный анализ (Multiple Regression) применяется для построения

уравнения регрессии для параметра Y в зависимости от факторов $X_1 \div X_k$. Модель может быть линейной и нелинейной. Наиболее простой, содержащей только линейные эффекты факторов, является линейная модель:

$$\hat{y} = b_o + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k , \qquad (1.1)$$

где $\hat{y}$ - прогнозируемое значение выходного параметра; b_o - свободный член; b_1, b_2, ..., b_k - коэффициенты регрессии; x_1, x_2, ..., x_k - возможные значения факторов X_1, X_2, ..., X_k; $b_1 x_1$, $b_2 x_2$, ..., $b_k x_k$ - линейные эффекты факторов.

Коэффициенты модели получают методом наименьших квадратов по исходной матрице наблюдений $n \times (k+1)$, где: n - число строк в матрице, равное числу наблюдаемых объектов, $k+1$ - число столбцов, равное числу независимых переменных (k факторов X_1, X_2, ..., X_k) и одной зависимой переменной (моделируемый параметр Y).

Значимость коэффициентов оценивают по *t*-критерию Стьюдента. При построении модели в ответственных случаях, например, для прогноза параметра Y, в модели сохраняют только значимые коэффициенты с доверительной вероятностью больше или равной 0,95 или с уровнем значимости $p \leq 0{,}05$. В поисковом исследовании с целью изучения характера изменения параметра Y при изменении факторов и степени влияния их на параметр, допускают сохранение в модели эффектов с коэффициентами, при уровне их значимости $p \leq 0{,}30$ (доверительной вероятностью равной или больше 0,70).

Стандартный алгоритм регрессионного анализа предусматривает расчет:

- числовых характеристик переменных;
- корреляционной матрицы;
- коэффициентов модели с оценками их значимости;

- результатов дисперсионного анализа модели и оценки коэффициентов множественной корреляции и детерминации, средней квадратичной ошибки прогноза параметра Y по модели;

- графика линии регрессии с указанием 95%-го доверительного интервала для прогноза значений параметра Y.

Вариант многофакторного регрессионного анализа с целью построения линейного уравнения регрессии дан в подглаве 1.2.

Сущность пошагового регрессионного анализа. Стандартный алгоритм многомерного регрессионного анализа обеспечивает получение коэффициентов модели для всех независимых переменных $X_1 \ldots X_k$. Исходя из уровней значимости, исследователь решает, какие коэффициенты должны быть включены в модель как значимые, достоверные. Для автоматического включения значимых эффектов в модель и исключения незначимых предлагается пошаговый регрессионный анализ в двух вариантах:

- Forward – поочередное включение в модель наиболее значимых эффектов;

- Backward – поочередное исключение из полной модели наименее значимых эффектов.

Отбор значимых эффектов реализуется по критерию F-Фишера.

В ответственных исследованиях для получения коэффициентов с уровнем значимости $p<0,05$ задаётся значение критерия $F=3 \div 4$. В поисковых исследованиях значение $F=1 \div 2$ обеспечивает включение в модель коэффициентов с уровнем значимости $p<0,30$.

Дисперсионный анализ и оценка эффективности модели. Дисперсионный анализ модели выполняется для оценки ее эффективности. Под эффективностью модели понимают её информативность и значимость (достоверность). Модель считают информативной, если ее коэффициент детерминации $R^2>0,5$; значимой, достоверной при уровне значимости по F - критерию $p \le 0,05$ (достоверности $\ge 0,95$).

Оценка степени влияния факторов на моделируемый параметр. Степень влияния факторов на параметр *Y* рассчитывается по величине стандартизованных коэффициентов регрессии *BETA* по формуле (1.2.):

$$K_j = \frac{100 \times BETA}{\sum_{(j)} |BETA_j|} \times R^2, \text{ в \%.} \qquad (1.2)$$

Прогноз по модели и оценка его точности и надежности. Прогноз среднеожидаемых значений параметра может быть дан по модели (1.1.) или графику линии регрессии. Точность и надежность прогнозируемого значения параметра оценивается 95%-м доверительным интервалом:

$$Y = \hat{y} \pm t_{95} \times m_{\hat{y}}, \qquad (1.3)$$

где $m_{\hat{y}}$ — среднеожидаемые значения параметра *Y*.

Достаточно большой доверительный интервал характерен для поискового исследования из-за приблизительной оценки значений факторов и параметра в матрице наблюдений [1-5].

Приведённые методы анализа сложных систем использовались авторами монографии в исследованиях данных биомедицинской информации.

1.2. Ферментная диагностика инфаркта миокарда с использованием статистических методов

Заболевания сердечно-сосудистой системы, прежде всего хроническая сердечная недостаточность (ХСН), ишемическая болезнь сердца и артериальная гипертензия являются основной проблемой здравоохранения экономически развитых стран и будет таковой в течение ближайших 50 лет. В структуре общей смертности эти заболевания составляют более половины всех случаев. Прямые затраты только на лечение ХСН на примере США составляют до 40 миллиардов долларов в год.

В арсенале современной медицины есть немало эффективных диагностических средств, однако часть из них инвазивна, опасна для пациента, а часть достаточно сложна в эксплуатации и имеет крайне

высокую цену. Большинство этих средств доступны лишь многопрофильным госпиталям и крупным (как правило, коммерческим) медицинским центрам и недоступны основной части больных. Одной из причин сложившейся ситуации являются невозможность и нежелание людей проводить раннюю диагностику – тратить свои время и денежные средства на проведение длительных и весьма дорогостоящих обследований.

Учитывая масштабы надвигающейся «неинфекционной пандемии» сердечно-сосудистых заболеваний современная медицина, особенно на уровне её первичного звена (терапевтический участок), должна быть вооружена простыми, безопасными для пациента, эффективными и надежными методиками своевременного, по возможности, максимально раннего выявления наиболее распространенных форм сердечно-сосудистой патологии. Одной из таких методик явилась ферментная диагностика, основой которой явилось обнаружение повышения содержания глутаминощавелевоуксусной трансаминазы (ГЩТ) при остром периоде инфаркта миокарда

При ферментной диагностике основное внимание уделяется изучению изменения следующих показателей: АсАТ (аспартатаминотрансфераза), АлАТ (аланинаминотрансфераза) и ЛДГ (лактатдегидрогеназа).

В типичных случаях острого инфаркта миокарда активность общей сывороточной ЛДГ увеличивается через 24-48 ч, достигает максимального уровня, превышая за 3-6 дней норму в 2-10 раз, и нормализуется через 8-14 дней. Наиболее высокая активность ЛДГ обнаружена у больных инфарктом миокарда, впоследствии умерших от разрыва сердца, в связи с чем предлагается расценивать значительное увеличение активности ЛДГ как фактор риска возникновения аневризмы и разрыва сердца. Активность общей ЛДГ при инфаркте миокарда повышается главным образом за счёт термостабильной фракции, которая возрастает в зависимости от обширности некроза. Наиболее выраженное и длительное усиление активности термостабильной фракции ЛДГ отмечалось у больных трансмуральным инфарктом миокарда. При этом максимальные её величины в 1,5 раза больше по сравнению с нормой [6].

Диагностические мероприятия при инфаркте миокарда могут быть разделены на две группы: собственно диагностика инфаркта миокарда и диагностика величины поражения мышцы сердца и последующих её изменений. К первой группе диагностических мероприятий относятся электрокардиография и ферментная диагностика, ко второй - прекардиальное картирование сердца с использованием множественных отведений для регистрации прекардиальных или эндокриальных ЭКГ, изменение активности изоферментов, КФК в крови и радиоизотопная визуализация острого инфаркта миокарда. С помощью последних трёх основных методов удалось не только количественно определить размер пораженного участка сердца, но и динамику процесса при острой коронарной катастрофе. Особенно хорошие результаты могут быть получены при их комплексном использовании.

В целом патогенез острого инфаркта миокарда предполагает следующие этапы развития:

1) механизмы формирования острой коронарной недостаточности;

2) метаболические и патоморфологические механизмы развития клеточного повреждения в зоне острой ишемии и будущего инфаркта миокарда;

3) активация нейрогуморальных факторов, участвующих в формировании ответно-адаптивных реакций на уровне целостного организма.

Авторами монографии получена корреляционная матрица (табл.1), содержащая основные показатели ферментной диагностики пациентов при инфаркте миокарда [7].

Таблица I

Показатели ферментной диагностики инфаркта миокарда

№	АсАТ	АлАТ	ЛДГ	Y, %
1	237	150	700	82
2	205	189	720	85
3	300	173	725	90
4	174	110	830	83
5	300	189	745	91
6	500	205	900	96
7	480	150	790	92
8	800	300	1050	94
9	450	284	895	97
10	322	126	724	89

Для обработки данных использовался регрессионный анализ в пакете STATISTICA 6.0 [8]. В многофакторной регрессионной модели (Relate\Multiple regression) зависимой переменной является Y – степень прогрессирования заболевания в процентах, а независимыми переменными (факторами) – АсАТ, АлАТ и ЛДГ. На рис.1.1 представлена таблица экспериментальных данных.

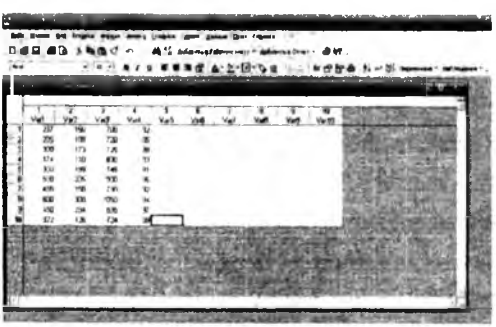


Рис.1.1 Экспериментальные данные

После процедуры выбора зависимых и независимых переменных и нажатия кнопки ОК (рис.1.2) STATISTICA выдаёт результаты множественной регрессии (рис.1.3).

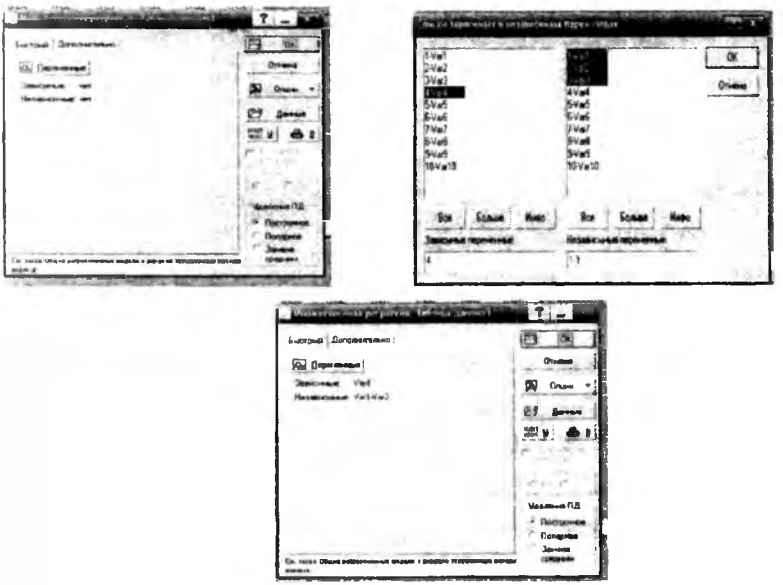


Рис.1.2. Процедура выбора переменных

Рис.1.3. Результаты множественной регрессии

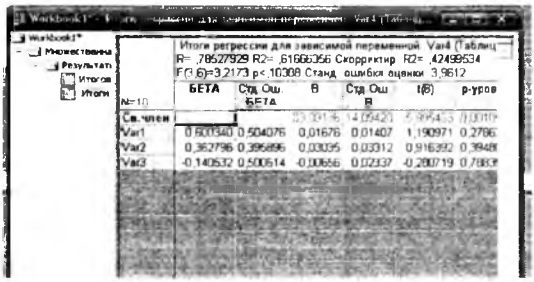


Рис.1.4. Итоговая таблица регрессии

На рис. 1.4 дан дисперсионный анализ модели, из которого следует, что модель достоверна (p=0,1). Из рис. 1.4 видно, что модель информативна, поскольку R^2=0,616 (больше 0,5), а стандартная ошибка прогноза возможных значений параметра S_0=14,094 и среднеожидаемых значений параметра

$$m_y = \frac{S_o}{\sqrt{n}} = \frac{14,094}{\sqrt{10}} = 3,96.$$

Такую модель можно применять для решения задач исследования.

Итоговая модель ферментной диагностики инфаркта миокарда в виде линейного уравнения регрессии имеет вид:

Y=83,092+ 0,016АсАТ+0,031АлАТ-0,006ЛДГ.

Найденная зависимость адекватно описывает эксперимент и позволяет «прогнозировать» значения зависимых переменных, то есть регрессионный анализ, помимо того что он позволяет количественно описывать зависимость между переменными, дает возможность прогнозировать значения зависимых переменных, – подставляя в найденную формулу значения независимых переменных, можно получать прогноз значений зависимых. При этом следует помнить, что построенная модель «локальна», то есть получена для некоторых вполне конкретных значений переменных. Экстраполяция результатов модели на более широкие области значений переменных может привести к ошибочным выводам.

ГЛАВА 2 . НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ ДЛЯ АНАЛИЗА И ОБРАБОТКИ МЕДИЦИНСКИХ ДАННЫХ

Современные информационные технологии позволяют выйти на качественно новый уровень представления течения заболевания, а именно визуально, на основе соответствующих математических моделей, пространственно смоделировать типовое развитие патологического процесса при конкретном заболевании. При этом особое место занимают так называемые самообучающиеся интеллектуальные системы (СИС). Они основаны на методах автоматической классификации ситуаций из реальной медицинской практики или на методах обучения на примерах. Наиболее яркий пример СИС – искусственные нейронные сети.

Искусственные нейронные сети (ИНС; artificial neural networks) представляют собой нелинейную систему, позволяющую классифицировать данные гораздо лучше, чем обычно используемые линейные методы. В приложении к медицинской диагностике ИНС дают возможность значительно повысить специфичность метода, не снижая его чувствительность.

Наиболее важным отличием ИНС от остальных методов прогнозирования является возможность конструирования самим врачом-специалистом, который может передать нейронной сети свой индивидуальный опыт и опыт своих коллег или обучать сеть на реальных данных, полученных путем наблюдений. Нейронные сети способны принимать решения, основываясь на выявляемых ими скрытых закономерностях в многомерных данных. Положительное отличительное свойство ИНС состоит в том, что они не программируются, т.е. не используют никаких правил вывода для постановки диагноза, а обучаются делать это на примерах. В ряде случаев ИНС могут демонстрировать удивительные свойства, присущие мозгу человека, в том числе отыскивать закономерности в запутанных данных.

Другое, не менее важное, свойство нейронной сети состоит в способности к обучению и обобщению полученных знаний. Сеть обладает чертами, так называемого искусственного интеллекта. Натренированная на ограниченном множестве обучающих выборок, она обобщает накопленную информацию и вырабатывает ожидаемую реакцию применительно к данным, не обрабатывавшимся в процессе обучения.

Несмотря на значительное количество уже известных примеров практического приложения искусственных нейронных сетей, возможности их дальнейшего использования для обработки медицинских данных окончательно не исчерпаны, и можно предположить, что ИНС еще в течение многих лет будут одним из основных инструментов поддержки принятия решений в условиях отсутствия точных моделей реальных процессов и явлений.

2.1. Искусственные нейронные сети: строение, принципы работы и обучение

Искусственной нейронной сетью называют некоторое устройство, состоящее из большого числа простых параллельно работающих процессорных элементов – нейронов, соединенных адаптивными линиями передачи информации. Работа нейросети заключается в преобразовании входного вектора – исходные данные – в выходной вектор – полученное решение [9]. В настоящее время разработано несколько видов искусственных нейронных сетей. Наиболее популярными из них являются многослойный персептрон (МСП) и сеть Кохонена.

На рис. 2.1 представлена схема строения МСП. Входной слой (слой 1) – группа связей, по которым ИНС получает информацию из внешнего мира задачи [9]. Промежуточный слой (слой 2) – группа связей, которая обеспечивает возможность моделирования нелинейных функций. Выходной слой (слой k) – группа выходных связей, с которых снимаются выдаваемые сетью сигналы.

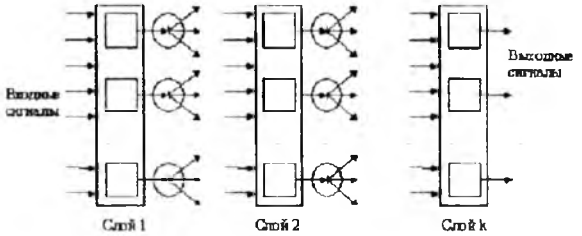


Рис. 2.1. Схема многослойного персептрона

Сеть подобной структуры может моделировать функцию практически любой степени сложности. Число слоев и число элементов в каждом слое определяется сложностью функции и подбирается эмпирически в ходе решения конкретной задачи. На практике создание подобной структуры осуществляется программными средствами. В настоящее время существует большое количество программных продуктов, эмулирующих нейросетевые алгоритмы обработки информации [10].

На рис 2.2. представлена схема строения самоорганизующейся карты признаков (сеть SOFM – Self- Organizing Feature Map) – сети Кохонена: *а* – входные элементы (предназначены только для того, чтобы распределять входные данные между выходными элементами); *б* – выходные (кластерные) элементы выступают в роли карты признаков пространства входных данных.

Выходы подобной сети должны трансформироваться в некоторую понятную форму, обусловленную процессом обучения. Например, при использовании пакета программных продуктов "Excel Neural Package" ответ сети представляет собой топологическую карту. Карта может быть представлена в линейной (рис. 2.3 а) или двухмерной (рис. 2.3 б) форме, которая задается пользователем. Ячейки топологической карты обозначаются латинскими буквами. В зависимости от конфигурации карты каждая ячейка может позиционироваться одним (только буквы, рис.

2.3 б) или двумя (буква + цифра, рис. 2.3 а) индексами. При получении ответа от сети на карте выделяются только заполненные ячейки. Так, на рис. 2.3 при конфигурации сети «4 на 4» (рис 2.2 а) на карте отмечено 7 ячеек, хотя теоретически возможное количество равно 16.

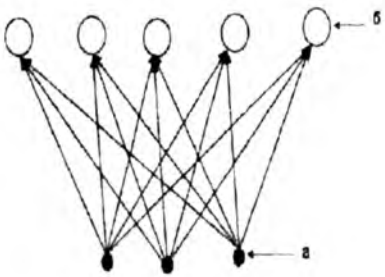


Рис. 2.2. Схема сети Кохонена

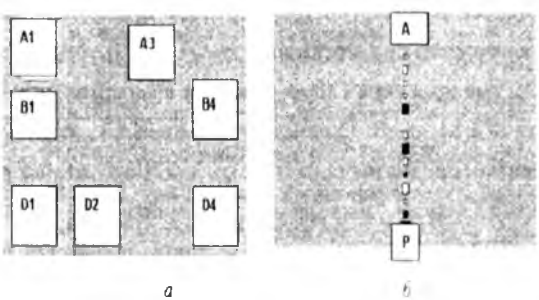


Рис. 2.3. Топологические карты с конфигурацией ячеек:
а - «4 на 4»; б - «16 на 1»

Сеть Кохонена, обучаясь на базе данных, способна построить двумерное отображение многомерных данных с минимально возможными искажениями (рис. 2.3). Такие самоорганизующиеся карты сохраняют локальную топологию данных - близость на карте подразумевает и близость в исходном пространстве базы данных. Сеть Кохонена может

распознавать кластеры в данных, а также устанавливать близость классов, т.е. реализует одно из свойств искусственных нейронных сетей - обобщение по подобию и одно из возможных применений таких сетей – разведочный анализ данных [11].

Ответственным моментом работы с ИНС является представление обучающих данных. Нейросети оперируют числовой информацией. Информация, на основании которой нейросеть должна давать ответ, может быть самого разнообразного вида: термины, описывающие какие-либо ситуации, числа различного вида и величины, графики, двух- и трехмерные изображения и т.д. Поэтому возникает необходимость корректного представления этой информации в виде чисел, сохраняющих смысл и внутренние взаимосвязи данных, т.е. приходится прибегать к элементарной числовой кодировке.

На рис. 2.4 показан механизм действия сети Кохонена при работе с базой данных. Выбор типа сети зависит от характера поставленной задачи. При решении задач прогнозирования и классификации чаще всего используются персептроны; при решении задач категоризации данных чаще всего применяется сеть Кохонена [12].

Эти две интеллектуальные задачи (классификацию и категоризацию данных) следует отличать друг от друга. Термин класс можно определить как совокупность предметов или понятий (образов), выделенных и сгруппированных по определенным признакам или правилам. Термин классификация подразумевает отнесение некоторого образа к классу, выполняемое по этим формальным правилам по совокупности признаков. Категория же (если отвлечься от специфического философского характера этого понятия) определяет лишь некоторые общие свойства образов и связи между ними.

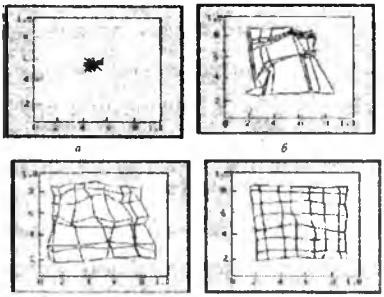


Рис. 2.4. Механизм действия сети Кохонена при работе с базой данных

Задача категоризации, т.е. определения отношения данного образа к некоторой категории, гораздо менее определена, чем задача отношения к классу. Границы различных категорий являются нечеткими, расплывчатыми, и обычно сама категория понимается не через формальное определение, а только в сравнении с другими категориями.

Границы классов, напротив, определены достаточно точно - образ относится к данному классу, если известно, что он обладает необходимым числом признаков, характерных для этого класса [13]. На рис.2.4:

а) неструктурированный массив базы данных (набор любых данных, в любой исследуемой области - чрезвычайно важная особенность нейросетевых программ, позволяющих решать любые вопросы классификации и категоризации);

б, в, г) этапы кластерного анализа данных.

Характер обучения зависит от выбранного типа сети. Для работы с многослойным персептроном используется «обучение с учителем». При этом способе обучения готовится набор данных, представляющих собой ряд наблюдений, для которых указаны значения входных и выходных переменных (условия задачи – ответ). Сеть учится устанавливать связь между ними [14]. На рис. 2.5 представлена схема обучения ИНС.

Рис. 2.5. Схема обучения искусственной нейронной сети

Важным моментом работы с искусственными нейронными сетями является представление обучающих данных. Нейросети оперируют числовой информацией. Данные, на основании которых нейросеть должна давать ответ, может быть самого разнообразного вида: термины, описывающие какие-либо ситуации, числа различного вида и величины, графики, двух- и трехмерные изображения и т.д.

Поэтому возникает необходимость корректного представления этой информации в виде чисел, сохраняющих смысл и внутренние взаимосвязи данных, т.е. приходится прибегать к элементарной числовой кодировке [14].

Некоторые особенности кодирования информации при проведении медико-биологических исследований были разработаны или усовершенствованы в ходе научных исследований. Обучение сети производится следующим образом: база данных (набор обучающих пар) делится на две неравные части; большую часть используют как обучающую, а меньшую как тестирующую базу; обучающая база вводится в нейросеть, сеть дает ответ; если ответы сети совпадают с экспертной оценкой – сеть обучена; если ошибка велика, то процесс обучения повторяется до тех пор, пока не будет получен результат, удовлетворяющий пользователя. При работе с сетью Кохонена используется «обучение без учителя». Процесс обучения, при котором на вход нейронной сети подаются данные, содержащие только значения

входных переменных. Такие алгоритмы предназначены для нахождения кластеров во входных данных [13]. Критерием достижения цели (обучение сети) считается результат тестирования набором примеров с известными ответами, не входящими в обучающую выборку [14].

2.2. Нейросетевые подходы обработки медицинской информации

Прежде чем приступить к работе с нейронной сетью, исследователю необходимо создать задачник или медицинскую базу данных (обучающую выборку). Такая база обычно формируется в виде таблицы и состоит из некоторого количества примеров, обладающих многими свойствами и особенностями (признаками) и, в зависимости от задачи, имеющих результат (результаты) вычисления по каждому примеру. В свою очередь, признаки могут носить различный характер и быть: существенными и второстепенными; качественными и количественными; первичными и вторичными и т.д. [15-18]. При этом для работы с нейронной сетью (опять же в зависимости от поставленной задачи) к выборке не предъявляется такое количество жестких требований, как, например, при работе с традиционными методами обработки информации. Структура таблицы экспериментальных данных приведена в табл.2.

Таблица 2

База данных

Объекты	Исходные признаки						Результаты		
	x_1	x_2	…	x_j	…	x_M	Y_1	…	Y_p
X^1	x_1^1	x_2^1	…	x_j^1	…	x_M^1	Y_1^1	…	Y_p^1
X^2	x_1^2	x_2^2	…	x_j^2	…	x_M^2	Y_1^2	…	Y_p^2
…	…	…	…	…	…	…	…	…	…
X^i	x_1^i	x_2^i	…	x_j^i	…	x_M^i	Y_1^i	…	Y_p^i
…	…	…	…	…	…	…	…	…	…
X^N	x_1^N	x_2^N	…	x_j^N	…	x_M^N	Y_1^N	…	Y_p^N

Такая таблица содержит N примеров, для характеристики которых используется некоторое количество признаков (равное M) и по заранее известному количеству результатов (равное P).

Для формирования такой таблицы необходимо обратить внимание на количество примеров (чем их больше, тем точнее результат вычисления); на информативную загрузку набора признаков: для начала работы лучше взять как можно больше признаков (впоследствии этот набор можно будет сократить), при этом их количество ограничивается лишь возможностями программного обеспечения.

Нейронная сеть способна обучаться решению задачи на основании такой обучающей выборки — "задачника", состоящего из набора пар "вход требуемый выход", и далее может решать примеры, не входящие в обучающую выборку [19, 20]. При этом процесс обучения представляет собой автоматический поиск закономерности между данными и заранее известным результатом в обучающей выборке. Примеры из предложенной выборки предъявляются нейросетевой модели, а нейроны, получая по входным связям сигналы — "условия примера", преобразуют их, несколько раз обмениваются преобразованными сигналами и выдают ответ (набор сигналов) — например, нозологии сердечной недостаточности — ишемическая болезнь сердца, стенокардия, инфаркт миокарда и т.д. [21, 22].

Обученная нейронная сеть автоматически записывается на диск компьютера как обыкновенный файл. В любой момент времени можно считать сеть с диска и продолжить обучение со старой или новой обучающей выборкой. Одна нейронная сеть обучается решать только одну задачу, однако может использовать для обучения различные обучающие выборки, которые могут различаться по количеству примеров, но должны соответствовать друг другу по числу обучающих параметров и их порядку и смыслу.

Тестирование выборки с заранее известными результатами позволяет проверить, правильно ли сеть определяет результаты для всех примеров и

насколько уверенно она это делает. Полученный результат по каждому примеру сравнивается с заранее известным, и если сеть обучилась полностью, то при тестировании той же самой выборки результаты будут определены правильно.

Кроме такого алгоритма тестирования можно также протестировать выборки с примерами, которые не участвовали в обучении нейронной сети. В этом случае можно спрогнозировать неизвестный результат или выявить недостатки в обучающей выборке, в структуре сети и в постановке задачи.

Изменяя в различных направлениях значения параметров примера и повторяя его тестирования, можно определить, как и насколько изменяется значение результата. Это позволяет создать почву для составления рекомендаций по изменению результата в реальной постановке диагноза.

В процессе обучения нейронная сеть способна оценивать влияние каждого из обучающих параметров на принятие решения (то есть определить степень важности отдельных признаков). Эта способность позволяет выявить информативность каждого признака по сравнению с остальными и оценить значимость признаков [23-26].

Применение этих алгоритмов и вычисленные на их основе показатели значимости признаков позволяют анализировать вектор входных сигналов на избыточность. То есть, отбрасывая сигналы, показатель которых имеет маленькое значение для вычисления ответа, обучаем сеть без него и смотрим изменение ошибки. В том случае, когда значение ошибки укладывается в заданный интервал, таким признаком можно пренебречь (исключить из выборки), а когда результат вычисляется более точно, можно говорить о зашумлении выборки этим признаком.

Данная возможность нейронных сетей (оценка значимости признаков) позволяет решить еще одну проблему: облегчение процесса создания вопросника для решения конкретной задачи. То есть для начала можно сформировать обучающую выборку с большим количеством признаков, а

затем нейроимитатор способен сам расклассифицировать их по важности и сократить список входных сигналов.

С помощью нейронных сетей можно автоматически минимизировать число обучающих признаков. При этом сокращение множества параметров и входных сигналов нейросети может преследовать несколько целей:

• упрощение специализированных технических устройств;

• сокращение объема используемой памяти и увеличение быстродействия нейросети при решении задачи;

• удешевление и упрощение сбора данных за счет отбрасывания наименее значимых;

• облегчение явной вербальной интерпретации процесса и результатов обработки данных.

Алгоритмы минимизации пространства медицинских признаков-симптомов заболеваний рассмотрены в [27].

Одними из исследователей, занимающихся теоретическими вопросами в области нейроинформатики и разрабатывающих программное обеспечение для решения прикладных задач в этой области, являются члены группы НейроКомп Красноярского института вычислительного моделирования СО РАН (Россия). Ими был разработан пакет MultiNeuron, который состоит из трех программ и является программным инструментарием, позволяющим создавать экспертные системы, и NeuroPro.

Интеллектуальные системы на основе искусственных нейронных сетей позволяют с успехом решать проблемы распознавания образов, выполнения прогнозов, оптимизации, ассоциативной памяти и управления. Они предпочтительны там, где имеется много входных данных, в которых скрыты закономерности. В этом случае можно почти автоматически учесть различные нелинейные взаимодействия между показателями-признаками, характеризующими такие данные. Это особенно важно в системах обработки информации, в частности, для ее предварительного анализа или

отбора, выявления "выпадающих фактов" (условно незначимых признаков) или грубых ошибок врача, принимающего решения.

Таким образом, имея достаточно простой набор элементов структуры, нейронные сети способны решать разнообразные медицинские задачи: а) управление в реальном времени; б) распознавание образов; в) предсказание; г) оптимизация; д) задачи обработки биомедицинских сигналов при наличии больших шумов и т.д. Можно сказать, что нейронные сети универсальный инструмент, с точки зрения медицинского исследования они являются достаточно действенными и зависят лишь от производительности компьютерных систем, они обеспечивают достаточно высокую устойчивость к ошибкам, на сегодняшний день они являются дешевым продуктом. Методы нейронных сетей могут использоваться или служить прекрасным дополнением к традиционным методам статистического анализа. Таким образом, применение технологии нейронных сетей выгодно, а разработка методов нейросетевого моделирования и анализа информации является актуальной задачей.

2.3. Диагностические медицинские программы, основанные на нейроприложениях

Анализ применения нейросетевых технологий в медицине показывает, что практически отсутствуют какие-либо методологии разработки нейросетевых медицинских систем, о чем свидетельствует как отсутствие работ такого профиля, так и огромное разнообразие подходов к нейросетевым алгоритмам обучения и архитектурам нейронных сетей. Всё это служит подтверждением того, что медицинская нейроинформатика как наука находится еще, в основном, на стадии накопления фактического материала.

Рассмотрим несколько наиболее интересных нейросетевых приложений для биологии и медицины, созданных различными авторами и школами.

Наибольший интерес для практического здравоохранения представляют системы для диагностики и дифференциальной диагностики заболеваний. При этом для принятия решений могут использоваться самые разнообразные данные - анамнез, клинический осмотр (создаются экспертные системы диагностики, ограничивающиеся только этим набором), результаты лабораторных тестов и сложных функциональных методов. Список областей медицины, в которых начали применяться новые технологии, чрезвычайно обширен и продолжает расти.

Одним из наиболее интенсивно развиваемых направлений является применение нейросетей в кардиологии.

В Италии разработана экспертная система для диагностики и лечения артериальной гипертонии [28]. Система включает в себя три нейросетевых модуля, причем ответы одних являются входными данными для других. В начале исследования больному проводят измерение систолического и диастолического давления каждые полчаса в течение суток. Данные за каждый час усредняются. Таким образом, образуется массив из 48 величин артериального давления (по 24 для систолического и диастолического). После этого первый модуль, состоящий из двух трехслойных нейросетей (в каждой из которых 2 входных, 4 "скрытых" и 24 выходных нейрона), на основании данных о поле и возрасте больного рассчитывает аналогичные "должные" величины и сравнивают их с реальными. Параллельно второй модуль (двухслойная нейросеть с 17 входными и 4 выходными нейронами) на основании клинических данных (симптоматика, анамнез) рассчитывает возможные сочетания гипотензивных лекарственных средств, которые могут быть использованы для лечения обследуемого больного.

Данные, снятые с выходов обоих модулей, вместе с клиническими данными подаются на вход последнего, третьего модуля (6-слойная нейросеть). Этот модуль оперирует 4 группами гипотензивных препаратов (диуретики, бетаадреноблокаторы, ингибиторы ангиотензина, блокаторы кальциевых каналов). Цель - назначить суточный (почасовой) график приема больным лекарств каждой (если требуется) из 4 групп. Поэтому

этот модуль имеет 96 выходных нейронов (4 препарата х 24 часа). С каждого выходного нейрона снимается доза, соответствующая одному препарату, назначаемому на данный час суток. Естественно, что в реальной ситуации большинство выходных данных равны нулю. Таким образом, создается оптимальная для пациента схема лечения гипертонии.

Нужно отметить, что система учитывает некоторые особенности приема препаратов больными, например, затруднение приема препаратов ночью (назначает ночной прием только в крайних случаях), запрет на назначение мочегонных лекарств на ночь. Отличительной чертой системы является возможность пользователя (врача) передавать нейронной сети свой опыт. Для этого создателями программы предусмотрен специальный блок, который выводит на экран компьютера суточные кривые артериального давления и предлагает врачу ввести в компьютер суточную схему приема гипотензивных препаратов в необходимых, по его мнению, дозах. Введенный пример помещается в базу данных. В любое время можно инициировать доучивание нейронных сетей с новыми примерами.

В одной из работ [29] приводится метод выявления атеросклеротических бляшек в артериях. Для этого применяется нейросеть, интерпретирующая флюоресцентные спектры, получаемые при исследовании тканей с помощью лазера. Аналогичным образом проводится диагностика заболеваний периферических сосудов, например, определение форм артериита.

Проводится комплекс исследований по использованию нейросетей для диагностики инфаркта миокарда [30-32]. Автор приводит данные по чувствительности (77,7%) и специфичности (97,2%) нейросетевого теста. В работе, кроме того, с помощью нейронной сети устанавливали диагностическую значимость клинических параметров при диагностике инфаркта миокарда.

Нейросетевой анализ акустических сигналов позволяет проводить диагностику клапанных шумов сердца [33] и оценивать систолическую и

диастолическую фазы сердечного сокращения с постановкой предварительного диагноза [34].

Нейросети используются терапевтами для диагностики заболеваний печени по лабораторным данным исследования функций печени [35]; дифференциальной диагностики заболеваний печени [36] и желчного пузыря по УЗИ [37].

Нейропрограммы могут с успехом работать с медицинскими данными, относящимися к субъективным категориям, например, в психиатрии [38]. Оценка субъективных данных дает возможность распознавания психических симптомов и диагностики [39] и изучения некоторых психиатрических симптомокомплексов [40].

Актуальная проблема диагностики злокачественных новообразований, возможно, получит новый уровень осмысления с началом применения нейроалгоритмов. Так, в работе показана 80%-я точность ранней диагностики меланом кожи – весьма распространённого злокачественного заболевания [41].

Одним из серьезных направлений применения нейронных сетей является интерпретация медицинских данных. В последние годы идет бурное развитие новых средств диагностики и лечения. При этом наблюдается "вторая волна" изучения и использования древних, старинных методов, и, наоборот, применение последних технических новшеств. Нередко и те и другие методы при использовании предоставляют врачу массу самых разнообразных данных. При этом встает проблема их грамотной и корректной интерпретации. Поиск глубинных закономерностей между получаемыми данными и патологическими процессами начинает отставать от разработки все новых и новых методов, поэтому применение для этой цели нейросетей может оказаться чрезвычайно выгодным.

Распознавание и интерпретация данных требует огромного опыта врача, что практически невозможно в современных условиях. Нейросеть была применена для "узкой" диагностики только по одной из точек пульса,

позволяющей оценивать состояние левой почки [42]. Пульс считывается специальным датчиком, совмещенным с микрофоном. Полученная пульсовая кривая (сфигмограмма) передается в компьютер. Вначале программа анализирует несколько пульсовых волн и выстраивает "среднюю" волну. После этого по 5 точкам этой волны нейронная сеть оценивает состояние левой почки.

Классической проблемой в кардиологии является интерпретация электрокардиограмм, требующая значительного опыта врача. Сотрудники Университета Глазго (Великобритания) ведут исследования по применению нейросетей для ЭКГ-диагностики инфарктов миокарда [43-46]. Входными данными для сетей являются избранные параметры 12-канальной электрокардиограммы и 12-канальной векторкардиограммы (длины зубцов, расстояния между зубцами). Исследователи обучили огромное количество нейросетей (167 сетей для диагностики инфаркта миокарда передней стенки и 139 сетей для инфаркта нижней стенки) на массиве данных из 360 электрокардиограмм. Обученные сети затем тестировали отдельную выборку с заранее известными ответами (493 случая). Одновременно для получения отдельной серии ответов на тестируемой выборке был использован логический метод (с заранее заданным алгоритмом). Затем сравнивались результаты тестирования выборки лучшими нейросетями и с помощью логического алгоритма. Сравнение показало, что во многих случаях чувствительность и специфичность нейросетевого теста оказались выше, чем у логического метода. Авторы делают справедливый вывод, что в случаях, когда логический алгоритм решения задачи все-таки можно выстроить, разумно комбинировать в экспертных системах оба подхода.

Интерпретация ЭКГ с помощью нейросетей была применена для диагностики злокачественных желудочковых аритмий [47]. Трехслойная сеть с 230 входными синапсами была обучена на 190 пациентов (114 с хронической сердечной недостаточностью и 34 с дилятационной кардиопатией) различать наличие (у 71 пациента) и отсутствие (у 119

пациентов) желудочковой тахикардии. Результаты тестирования сравнивались с логическим методом интерпретации данных. Показано, что нейросетевой тест обладает большей чувствительностью (73% по сравнению с 70 для логического метода) и специфичностью (83 и 59%).

Интересная работа описывает моделирование применения нейросетей для работы электрокардиостимуляторов (искусственных водителей ритма) [48]. Выпускаемые за рубежом электрокардиостимуляторы задают ритм не жестко, а в зависимости от исходного ритма, генерируемого синусовым узлом сердца. Например, если синусовый узел при какой-либо патологии генерирует недостаточное количество импульсов, водитель ритма компенсирует ритм. Таким образом, электрокардиостимулятор представляет собой систему вход-преобразование-выход, где входом является ритм синусового узла, выходом – собственный ритм электрокардиостимулятора, а преобразование осуществляется по заданному логическому алгоритму. Авторы смоделировали замену логического преобразователя нейронной сетью, так как взаимоотношения между генерацией импульсов в синусовом узле и требуемым ритмом не линейны и применяемые алгоритмы на практике не всегда эффективны. Нейросеть, обученная на 27 здоровых людях в ситуациях с различной физической нагрузкой, показала гораздо лучшую способность задавать ритм, чем логический алгоритм, применяющийся в электрокардиостимуляторе.

Одной из самых сложных задач для нейросетей в практической медицине является обработка и распознавание сложных образов, например рентгенограмм. Разработана экспертная система [49] интерпретации рентгенограмм груди у новорожденных с выбором одного и более диагнозов из 12.

Созданы нейросетевые экспертные системы для классификации опухолей молочной железы (определение – доброкачественная опухоль или злокачественная) по данным маммографии (сканограмма молочной железы) [50, 51]. По данным, которые приводят авторы, точность такого

вывода до применения нейросети составляла не более 75%. При тестировании системы нейросеть, анализирующая сканограмму, давала правильный ответ в 100% случаев. При тестировании изображение, получаемое в результате использования метода, представляется в виде матрицы точек размером 1024x1024 пиксела с 10-битовой шкалой яркости. Изображение подается на нейросеть, имеющую 2 входных, 80 "скрытых" и 2 выходных нейрона. При этом один из выходных нейронов "отвечает" за доброкачественную опухоль, другой за злокачественную. Диагноз определяется в зависимости от выходного нейрона, выдавшего больший по величине ответ. Столь высокий процент правильности распознавания, возможно, случаен и объясняется недостаточным количеством примеров, использовавшихся при обучении и тестировании нейросети (по 10 примеров). Однако даже при такой малой обучающей выборке нейросеть выигрывала по сравнению с традиционным методом интерпретации сканограммы.

Несколько работ посвящены нейросетевой обработке лабораторных анализов и тестов. Приводится нейросетевой метод интерпретации лабораторных данных биохимического анализа крови [52]. В работе показаны преимущества нейронных сетей в сравнении с линейным дискриминантным анализом, которым параллельно обрабатывались данные. Описываются обработка нейросетями цитогенетических данных [53] и интерпретация данных иммунного анализа [54].

Особое место среди нейросетевых экспертных систем занимают прогностические модели, применяемые, например, для прогнозирования исходов заболеваний.

В 1990 году американская фирма "Апачи Медикл Системз Инк." установила в реанимационном отделении одной из больниц штата Мичиган экспертную систему "Апачи - III" [55]. Ее цель - прогнозирование исхода заболевания у больных, находящихся в тяжелом состоянии. Для ноза в компьютер необходимо ввести 27 параметров больного: рвичный диагноз, симптомы, степень утраты сознания, наличие или

отсутствие СПИД и других заболеваний. После этого система выдает вероятность выживания больного в диапазоне от 0 до 100 процентов. Ценность применения системы заключается в том, что она позволяет очень быстро оценить динамику изменения состояния больного, незаметную "на глаз".

Например, можно получить ответ у системы до и после введения какого-либо лекарства и, сравнив ответы, посмотреть, будет ли наблюдаться эффект от терапии. Без программы же изменение состояния иногда не удается обнаружить в течение нескольких дней. Тестирование показало, что 95% прогнозов, которые делает программа, сбываются с точностью до 3%, что значительно точнее, чем у лучших врачей. Необходимо отметить, что система была обучена на данных, взятых из историй болезней 17448 пациентов, лечившихся в 40 больницах штата в 1989 году. Очевидно, что если качество работы системы обеспечивается таким большим объемом выборки, возможности перенастройки системы не слишком велики. Концепция авторов, создавших эту систему, заключается в как можно большем охвате различных примеров и вариантов (сбор данных в 40 больницах), а не в возможности индивидуализации системы к конкретной клинике. Поэтому данная система не способна к подучиванию в процессе работы, опыт "зашит" в нее жестко

Это может быть существенным недостатком при установке программы в регионы, резко отличающиеся по социально-географическим условиям от тех, где проводилось обучение. Кроме того, огромный массив примеров для обучения повышает стоимость программы.

В настоящее время нейросетевые технологии приобретают все большую популярность среди широкого круга аналитиков и исследователей. Несомненно, это связано с бурным развитием программного обеспечения и компьютерных технологий и, в первую очередь, с появлением доступных программных средств, имеющих широкий набор возможностей для нейросетевого моделирования. Среди

лидеров в этой области – STATISTICA Neural Networks, которые являются удобным и эффективным инструментом создания, обучения, исследования и автоматического конструирования нейронных сетей. Важнейшим достоинством пакета SNN является открытый API-интерфейс, т.е. набор библиотек с исчерпывающим набором функций, которые можно использовать для включения нейросетевых модулей в собственные приложения [56].

Применение нейронной сети для диагностики объясняется необходимостью идентифицирования функции, позволяющей выделять определенное заболевание. Существуют следующие методы определения данной функции:

1. Приближение полиномом (оценка состояния осуществляется по коэффициентам полинома путем вывода соответствия между значениями коэффициентов полинома и реакциями системы).

2. Приближение экспоненциальной функцией.

3. Обучение системы на основе нейронной сети (НС) позволяет с высокой точностью определить вид функции для диагностического анализа конкретного заболевания [56].

2.4. Информационные технологии Data Mining

Примером другой перспективной технологии обработки и обобщения больших объемов информации для решения задач классификации и прогнозирования в медицинских исследованиях является так называемая технология анализа и добычи данных Data Mining [56]. Методы и инструментальные средства анализа и добычи данных представляют собой дальнейшее развитие таких известных статистических инструментов разведочного анализа, как метод главных и метод независимых компонент, факторный анализ, множественная регрессия, редуцирование пространства признаков с использованием метода многомерного шкалирования, кластерного анализа и распознавания образов и др.

Программно реализованные и снабженные удобным пользовательским интерфейсом, а также поддержанные гибкими алгоритмами визуализации многомерных данных, средства Data Mining позволяют проводить соответствующие исследования даже начинающему пользователю. В арсенал методов кластерного анализа и распознавания образов систем Data Mining обычно входят метод опорных векторов (Support Vector Machine, или SVM), метод деревьев решений (decision trees), метод «ближайшего соседа» в пространстве признаков, байесовская классификация и др. Среди указанной группы методов классификации и распознавания наиболее интересным и гибким представляется метод опорных векторов (МОВ).

Решение задач классификации и прогнозирования очень важно для медицинских работников. Технологии StatSoft позволяют успешно решать эти задачи и эффективно применять на практике разнообразные методы анализа и добычи данных:

- данные имеют неограниченный объем;
- данные разнородны: количественные, качественные, текстовые;
- результаты должны быть конкретны и понятны;
- инструменты для обработки сырых данных должны быть просты в использовании.

Образ – это все, что может называться информацией, т.е. что-то, имеющее некоторые характеризующие его признаки:

- любые базы данных, состоящие из колонок цифр и строчек;
- любое оцифрованное изображение.

Технология Data Mining предоставляет возможность распознавать образы с применением метода статистически взвешенных синдромов:

1) сначала ставятся границы градации (не обязательно одна) по одному из признаков таким образом, чтобы с одной стороны границы преобладали значения одной из групп, а с другой стороны было больше значений второй группы. То же делается для второго признака, и для всех признаков, участвующих в обучении. Благодаря современной технике число признаков может быть почти не ограничено;

2) далее из всех признаков оставляют только те, которые наиболее информативны с точки зрения отделения одной группы от другой;

3) создается решающее правило, которое включает в себя набор наиболее информативных признаков с их границами градаций. По нему новый объект, не участвующий в обучении, можно распознать, т.е. отнести с некоторой вероятностью к одной из групп;

4) статистически взвешенное голосование проводится суммарно по всем базовым множествам;

5) в результате распознавания имеется одно число, находящееся между номерами групп. К какому номеру группы оно ближе, к той группе и будет относиться распознаваемый объект. Существует зона неопределенности. Если результат попадает в нее, то решение неопределенно.

6) самое главное – доказать, что различия между группами, найденные в результате распознавания, достоверно значимы. Для этого существует перестановочный тест, использующий метод Монте-Карло. Номера группы каждому объекту присваиваются произвольно и на скользящем контроле опять проводят обучение и распознавание. Так делается в автоматическом режиме тысячу раз. Если хорошее распознавание получается в 5 случаях из этой тысячи, то считается, что достоверность равна 0,005. Если таких случаев достаточно много, то скорее всего различия между исследуемыми группами нет.

Для медико-биологической информации характерны небольшие выборки, большое число параметров и наличие пропущенных значений в данных. Эти трудности, принципиальные для традиционных статистических методов, для Data Mining не страшны. Работать приходится именно с такими сложными данными. Причем данные могут быть и количественными, и качественными, непрерывными или дискретными. Главное, чтобы они имели вид таблицы, в которой один из столбцов является группирующим, т.е. содержит номера групп, к которым относятся каждый из объектов (данная строка - есть информация об одном

объекте, "запись"). Имеется в виду, что сравниваемые группы заранее известны. Это могут быть группы больных с различным исходом лечения или заболевания, группы экспериментальных данных и контроля и т.д. Обучение идет на данных с известным разделением на группы. После получения решающего правила можно любой предлагаемый новый объект, группа которого не известна, с некоторой вероятностью отнести к одной из групп, т.е. сделать для него прогноз или диагностику. Это и есть распознавание образов в действии.

С помощью данных методов распознавания образов были успешно решены многие задачи в медико-биологических исследованиях.

2.5. Диагностика сердечной патологии с помощью STATISTICA Data Miner

На основе данных, полученных по клиническим, функциональным и биохимическим исследованиям пациентов с кардиологическими заболеваниями, необходимо поставить дифференцированный диагноз. Были собраны данные по 43 пациентам, для каждого пациента известно до 38 характеристик (таких, как ЧСС, АД, коронарография, перфузионная сцинтиграфия миокарда и т.д.). Данные полные, в таблице нет пропущенных ячеек. Для постановки диагноза (норма, невыраженная ишемия, умеренная ишемия, выраженная ишемия) были использованы следующие методы: Neural Networks – нейронные сети; Machine Learning – процедуры обучения; C&RT Trees - деревья классификации и регрессии.

Результаты анализа по методам.

Нейронные сети. За 10 минут работы удалось построить и обучить сеть 100%-ной точности. При этом попутно были найдены и выделены всего 10 переменных из 40, которые полностью объясняли зависимую переменную – диагноз (рис.2.6).

Сеть – многослойный персептрон, всего 3 слоя (1 скрытый), 4 элемента на входном слое, 10 на промежуточном, 1 на выходном.

Процедуры обучения. Процедуры обучения (метод опорных векторов и Байесовские оценки) не показали в данной задаче хороших результатов. Процент правильной классификации составил 91%, причем диагноз ставился с точностью всего 60% (рис.2.7).

Деревья классификации и регрессии. Метод показал 100%-ый результат (рис.2.8), это связано с тем, что задача относительно несложная и число классов зависимой переменной невелико (4 класса). Как видно, деревья классификации с построением выделили небольшое количество переменных, полностью классифицирующих зависимую переменную – диагноз (рис.2.9).

Таким образом, с помощью методов Data Mining, реализованных в STATISTICA, удалось поставить диагноз с 100% точностью. Был выделен набор из 10 характеристик, которые определяют дифференцированный диагноз.

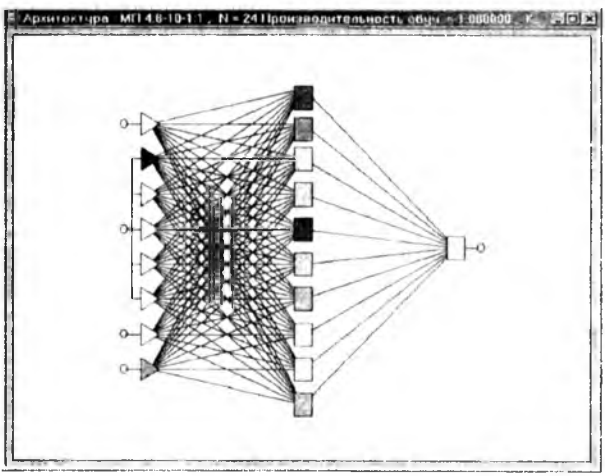


Рис. 2.6. Архитектура многослойного персептрона

Data: Classification summary (Support Vector Machine), P...

Classification summary (Support Vector Machine), PAT
SVM: Classification type 1 (C=10.000), Kernel: Radial
Number of support vectors= 32 (0 bounded)

Class Name	Total	Correct	Incorrect	Correct(%)	Incorrect(%)
complic	10	6	4	60.0000	40.00000
uncompl	33	33	0	100.0000	0.00000

Рис. 2.7. Итоги классификации методом опорных векторов

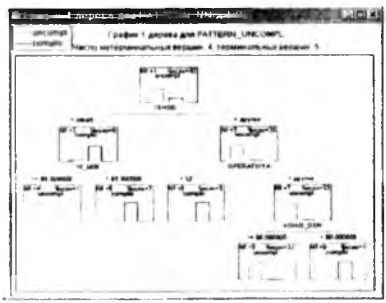


Рис.2.8. Дерево классификации

Данные: Матрица кла...

Матрица классифика
Зависимая перемен
Опции: Категориалы

	Класс uncompl	Класс complic
uncompl	33.00000	
complic		10.00000

Рис.2.9. Матрица ошибок классификации
с помощью деревьев классификации

2.6. STATISTICA Neural Networks в нейросетевых вычислениях

Использование нейронных сетей подразумевает гораздо большее, чем просто обработку данных нейросетевыми методами.

STATISTICA Neural Networks [57] предоставляет разнообразные функциональные возможности для работы с очень сложными задачами, включающие не только новейшие «архитектуры нейронных сетей» и «алгоритмы обучения», но также и новые подходы в «отборе входных данных» и «построении сети» (рис. 2.10). Кроме того, разработчики программного обеспечения предусмотрели следующую возможность после проведения заданных экспериментов в интерфейсе STATISTICA Neural Networks, нейросетевые анализы могут быть объединены в пользовательском приложении. Это достигается либо с помощью библиотеки COM-функций STATISTICA, которая полностью отражает все функциональные возможности программы, либо с помощью кода на языке C (C++, C#) или Visual Basic, который генерируется программой и помогает запустить полностью обученную нейронную сеть или сетевой ансамбль.

Модуль STATISTICA Neural Networks полностью интегрирован с системой STATISTICA, таким образом доступен огромный выбор инструментов редактирования (подготовки) данных для анализа (преобразования, условия выбора наблюдений, средства проверки данных и т.д.) (рис.2.11). Как и все анализы STATISTICA, программа может быть "присоединена" к удаленной базе данных с помощью инструментов обработки "на месте" или связана с активными данными, чтобы модели обучались или запускались (например, для вычисления предсказанных значений или классификации) автоматически каждый раз при изменении данных.

После того как данные подготовлены, необходимо решить, какие ...ченные следует использовать при работе с нейросетью. Чем больше

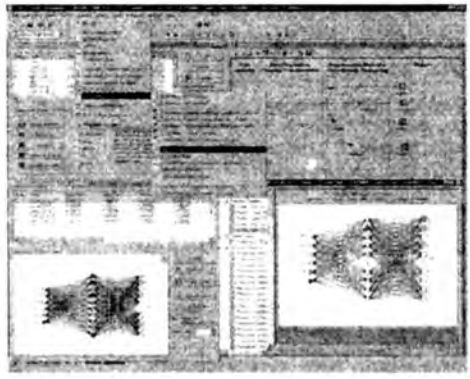


Рис. 2.10. Функциональные возможности STATISTICA Neural Networks

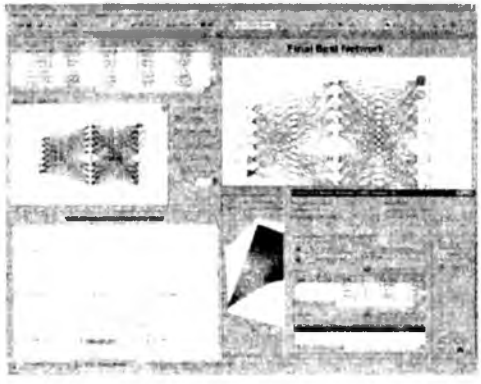


Рис. 2.11. Выбор инструментов STATISTICA Neural Networks

число переменных, тем сложнее будет нейронная сеть и, следовательно, потребуется больше памяти и времени на обучение, а также большее количество обучающих примеров (наблюдений). При недостаточном объеме данных и/или корреляциях между переменными исключительную важность во многих нейросетевых приложениях приобретают вопросы отбора значимых входных переменных и сжатия информации в меньшее число переменных.

В STATISTICA Neural Networks реализованы обратный и прямой алгоритмы пошагового выбора. Кроме того, нейрогенетический алгоритм отбора входных данных сочетает в себе возможности генетических алгоритмов и PNN/GRNN (PNN - вероятностные нейронные сети, GRNN - обобщенно-регрессионые нейронные сети) для автоматического поиска оптимальных комбинаций входных переменных, в том числе и в тех случаях, когда между ними имеются корреляции и нелинейные зависимости.

Почти мгновенная скорость обучения по PNN/GRNN-алгоритму не только дает возможность применить нейро-генетический алгоритм отбора входных данных, но и позволяет (пользуясь имеющимися в редакторе данных системы STATISTICA Neural Networks удобными средствами подавления незначимых переменных) в реальном времени проводить собственные эксперименты на чувствительность данных. STATISTICA Neural Networks содержит также встроенную систему «анализа главных компонент» (АГК и ассоциативные сети для "нелинейного АГК"), позволяющую понижать размерность исходных данных. В базовой системе STATISTICA доступно большое разнообразие статистических методов для понижения размерности данных.

Составной частью пакета STATISTICA Neural Networks является Intelligent Problem Solver – «мастер решений», который оценивает множество нейронных сетей различной архитектуры и сложности и выбирает сети наилучшей архитектуры для данной задачи.

Мастер способен строить сети для данных с независимыми наблюдениями (стандартные регрессионные сети, сети классификации или смешанные) равно как и сети, предназначенные для прогнозирования будущих значений некоторой переменной на основе уже имеющихся значений этой же переменной (сети с временными рядами).

Значительное время при создании нейронной сети уходит на выбор соответствующих переменных и оптимизацию архитектуры сети методом эвристического поиска. STATISTICA Neural Networks берет эту работу на

себя и автоматически проводит эвристический поиск. Эта процедура учитывает входную размерность, тип сети, размеры сети и требуемые выходные функции кодирования.

В ходе поиска можно задать число откликов, получаемых в процессе обучения. При установке режима максимальной детализации «мастер решений» будет отображать архитектуру и уровни качества для каждой тестируемой сети.

«Мастер решений» является чрезвычайно эффективным инструментом при использовании сложных техник, позволяя автоматически находить наилучшую архитектуру сети.

Автоматический конструктор сети можно также использовать в процессе разработки модели, когда модуль STATISTICA Neural Networks, совместно с другими модулями базовой системы STATISTICA, применяется для выявления наиболее значимых переменных (например, наилучших предикторов для последующего их включения и тестирования в какой-либо модели Nonlinear Estimation – нелинейного оценивания).

Успех экспериментов по поиску наилучшего типа и архитектуры сети существенным образом зависит от качества и скорости алгоритмов обучения сети. В системе STATISTICA Neural Networks реализованы самые лучшие на сегодняшний день обучающие алгоритмы.

Для обучения многослойных персептронов в системе STATISTICA Neural Networks реализован, прежде всего, метод обратного распространения - с меняющимися во времени скоростью обучения и коэффициентом инерции, перемешиванием наблюдений перед очередным шагом алгоритма и добавлением аддитивного шума для робастного обобщения. Помимо этого, в системе STATISTICA Neural Networks реализованы два быстрых алгоритма второго порядка - методы сопряженных градиентов и Левенберга-Маркара.

Последний представляет собой необычайно мощный современный алгоритм нелинейной оптимизации, и специалисты очень рекомендуют им пользоваться. В то же время область применения этого метода ограничена

случаями относительно небольших по размеру сетей с одним выходным нейроном, а для более громоздких задач в пакете STATISTICA Neural Networks имеется метод сопряженных градиентов. Как правило, и тот и другой алгоритм сходятся быстрее, чем метод обратного распространения, и при этом обычно выдают более подходящее решение.

Итеративный процесс обучения сети в системе STATISTICA Neural Networks сопровождается автоматическим отображением текущей ошибки обучения и вычисляемой независимо от нее ошибки на проверочном множестве, при этом показывается и график суммарной ошибки.

Обучение прерывается в любой момент нажатием кнопки. Кроме того, имеется возможность задать условия остановки, при выполнении которых обучение будет прервано; таким условием может быть, например, достижение определенного уровня ошибки, или стабильный рост проверочной ошибки на протяжении заданного числа проходов - "эпох" (что свидетельствует о так называемом переобучении сети). Если переобучение имеет место, это не должно заботить пользователя: STATISTICA Neural Networks автоматически запоминает экземпляр наилучшей сети, полученной в процессе обучения, и к этому варианту сети всегда можно обратиться, нажав соответствующую кнопку. После того, как обучение сети завершено, можно проверить качество ее работы на отдельном тестовом множестве.

В пакете STATISTICA Neural Networks реализован также целый ряд алгоритмов обучения для других сетей других архитектур. Параметры радиальных сплайнов и коэффициенты сглаживания у сетей на радиальной базисной функции и обобщенно-регрессионных сетей могут выбираться по таким алгоритмам, как: обучение Кохонена, подвыборка, метод *K*-средних, методы изотропии и ближайших соседей. Нейроны линейного выходного слоя у сетей на радиальной базисной функции, как и у линейных сетей, полностью оптимизируются методом сингулярного разложения (SVD).

2.7. Задачи регрессии в пакете STATISTICA Neural Networks

Задачи прогнозирования можно разбить на два основных класса: классификация и регрессия. В задачах классификации нужно бывает определить, к какому из нескольких заданных классов принадлежит данный входной набор. Примерами могут служить диагностика раковых заболеваний (опухоль, чисто). Во всех этих случаях, очевидно, на выходе требуется всего одна номинальная переменная. Чаще всего (как в этих примерах) задачи классификации бывают двузначными, хотя встречаются и задачи с несколькими возможными состояниями. В задачах регрессии требуется предсказать значение переменной, принимающей (как правило) непрерывные числовые значения. В таких случаях в качестве выходной требуется одна числовая переменная [58].

Нейронная сеть может решать одновременно несколько задач регрессии и/или классификации, однако обычно в каждый момент решается только одна задача. Таким образом, в большинстве случаев нейронная сеть будет иметь всего одну выходную переменную; в случае задач классификации со многими состояниями для этого может потребоваться несколько выходных элементов (этап пост-процессирования отвечает за преобразование информации из выходных элементов в выходную переменную).

В пакете STATISTICA Neural Networks для решения всех этих вопросов реализованы специальные средства пре- и пост-процессирования, которые позволяют привести сырые исходные данные в числовую форму, пригодную для обработки нейронной сетью, и преобразовать выход нейронной сети обратно в формат входных данных. Нейронная сеть служит "прослойкой" между пре- и пост-процессированием, и результат выдается в нужном виде (например, в задаче классификации выдается название выходного класса). Кроме того, в пакете STATISTICA Neural Networks пользователь может (если пожелает) получить прямой доступ к внутренним параметрам активации сети.

Согласно общепринятому в науке принципу, если более сложная модель не дает лучших результатов, чем более простая, то из них следует предпочесть вторую. В терминах аппроксимации отображений самой простой моделью будет линейная, в которой подгоночная функция определяется гиперплоскостью. В задаче классификации гиперплоскость размещается таким образом, чтобы она разделяла собой два класса (линейная дискриминантная функция); в задаче регрессии гиперплоскость должна проходить через заданные точки. Линейная модель обычно записывается с помощью матрицы $N \times N$ и вектора смещения размера N.

На языке нейронных сетей линейная модель представляется сетью с промежуточными слоями, которая в выходном слое содержит только линейные элементы (то есть элементы с линейной функцией активации). Веса соответствуют элементам матрицы, а пороги - компонентам вектора смещения. Во время работы сеть фактически умножает вектор входов на матрицу весов, а затем к полученному вектору прибавляет вектор смещения.

В пакете STATISTICA Neural Networks имеется возможность создать линейную сеть и обучить ее с помощью стандартного алгоритма линейной оптимизации, основанного на псевдообратных матрицах (SVD). Разумеется, метод линейной оптимизации реализован также в модуле Множественная регрессия системы STATISTICA; однако линейные сети пакета STATISTICA Neural Networks имеют то преимущество, что здесь можно в единой среде сравнивать такие сети с "настоящими" нейронными сетями.

Линейная сеть является хорошей точкой отсчета для оценки качества построенных пользователем нейронных сетей. Может оказаться так, что задачу, считавшуюся очень сложной, можно успешно решить не только нейронной сетью, но и простым линейным методом. Если же в задаче не так много обучающих данных, то, вероятно, просто нет оснований использовать более сложные модели.

В задачах регрессии целью является оценка значения числовой выходной переменной по значениям входных переменных. Задачи регрессии в пакете STATISTICA Neural Networks можно решать с помощью сетей следующих типов: многослойный персептрон, радиальная базисная функция, обобщенно-регрессионная сеть и линейная сеть. При этом выходные данные должны иметь стандартный числовой (не номинальный) тип.

Особую важность для регрессии имеют масштабирование (шкалирование) выходных значений и эффекты экстраполяции.

Нейронные сети наиболее часто используемых архитектур выдают выходные значения в некотором определенном диапазоне (например, на отрезке [0,1] в случае логистической функции активации). Для задач классификации это не создает трудностей. Однако для задач регрессии совершенно очевидно, что тут есть проблема, и некоторые ее детали оказываются весьма тонкими.

Для начала необходимо применить алгоритм масштабирования, чтобы выход сети имел "приемлемый" диапазон. Простейшей из масштабирующих функций пакета STATISTICA Neural Networks является минимаксная функция: она находит минимальное и максимальное значение переменной по обучающему множеству и выполняет линейное преобразование (с применением коэффициента масштаба и смещения), так чтобы значения лежали в нужном диапазоне (как правило, на отрезке [0.0,1.0]). Если эти действия применяются к числовой (непрерывной) выходной переменной, то есть гарантия, что все обучающие значения после преобразования попадут в область возможных выходных значений сети и следовательно, сеть может быть обучена. Кроме того, известно, что выходы сети должны находиться в определенных границах. Это обстоятельство можно считать достоинством или недостатком – здесь возникают вопросы экстраполяции.

На рис.2.12 иллюстрируется возможность оценки значения Y по значению X. Необходимо аппроксимировать кривую, проходящую через

заданные точки. Вероятно, вполне подходящей для этого покажется кривая, изображенная на графике – она (приблизительно) имеет нужную форму и позволяет оценить значение Y в случае, если входное значение лежит в интервале, который охватывается сплошной частью кривой - в этой области возможна интерполяция.

Но что делать, если входное значение расположено существенно правее имеющихся точек? В таких случаях возможны два подхода к оценке значения Y. Первый вариант – экстраполяция: подогнанная кривая продолжается вправо. Во втором варианте нет достаточной информации для осмысленной оценки этого значения, и потому в качестве оценки принимается среднее значение всех выходов (в отсутствие какой-либо информации это может оказаться лучшим выходом из положения).

Предположим, например, что используется многослойный персептрон (MLP). Применение минимакса по описанной выше схеме весьма ограничительно.

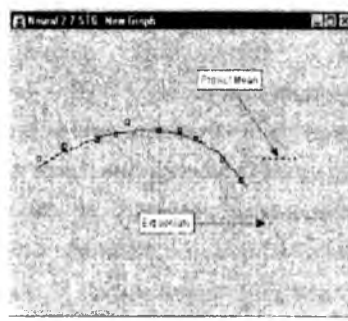


Рис. 2.12. Интерполяция функцией Y по значениям X

Во-первых, кривая не будет экстраполироваться, как бы близко ни оказалось расположение к обучающим данным. Во-вторых, оценка по среднему также не будет выполняться: вместо этого будет браться минимум или максимум смотря по тому, росла или убывала в этом месте

оцениваемая кривая. Чтобы избежать этих недостатков в MLP используется ряд приемов:

Во-первых, логистическую функцию активации в выходном слое можно заменить на линейную, которая не меняет уровня активации (N.B.: функции активации меняются только в выходном слое; в промежуточных слоях по-прежнему остаются логистические и гиперболические функции активации). Линейная функция активации не насыщается, и поэтому способна экстраполировать (при этом логистические функции предыдущих уровней все-таки предполагают насыщение на более высоких уровнях). Линейные функции активации в MLP могут вызвать определенные вычислительные трудности в алгоритме обратного распространения, поэтому при его использовании следует брать малые (менее 0.1) скорости обучения. Описанный подход пригоден для целей экстраполяции.

Во-вторых, можно изменить целевой диапазон минимаксной масштабирующей функции (например, сделать его [0.25,0.75]). В результате обучающие наблюдения будут отображаться в уровни, соответствующие средней части диапазона выходных значений. Интересно заметить, что если этот диапазон выбран маленьким и обе его границы находятся вблизи значения 0.5, то он будет соответствовать среднему участку сигмоидной кривой, на котором она "почти линейна", - тогда выполняется практически та же схема, что и в случае линейного выходного слоя. Такая сеть сможет выполнять экстраполяцию в определенных пределах, а затем будет насыщаться. Все это можно хорошо себе представить следующим образом: экстраполяция допустима в определенных границах, а вне их она будет пресекаться.

Если применяется первый подход и в выходном слое помещены линейные элементы, то может получиться так, что вообще нет необходимости использовать алгоритм масштабирования, поскольку элементы и без масштабирования могут выдавать любой уровень выходных сигналов. В пакете STATISTICA Neural Networks имеется возможность для большей эффективности вообще отключить всё

масштабирование. Однако на практике полный отказ от масштабирования приводит к трудностям в алгоритмах обучения. Действительно, в этом случае разные веса сети работают в сильно различающихся масштабах, это усложняет начальную инициализацию весов и (частично) обучение. Поэтому не рекомендуется отключать масштабирование, за исключением тех случаев, когда диапазон выходных значений очень мал и расположен вблизи нуля. Это же соображение говорит в пользу масштабирования и при препроцессировании в MLP-сетях (при котором, в принципе, веса первого промежуточного слоя можно легко корректировать, добиваясь этим любого нужного масштабирования).

Сети, в которых используются радиальные элементы (RBF и GRNN), работают совершенно иначе, чем сети MLP. Радиальные сети по самой своей природе неспособны к экстраполяции. Чем дальше входной пример расположен от точек, соответствующих радиальным элементам, тем меньше становятся уровни активации радиальных элементов и (в конце концов) тем меньше будет выходной сигнал сети. Входной пример, расположенный далеко от центров радиальных элементов, даст нулевой выходной сигнал. Стремление сети не экстраполировать данные можно считать достоинством (это зависит от предметной области), однако убывание выходного сигнала (на первый взгляд) достоинством не является. Если необходимо избежать экстраполяции, то для входных точек, отличающихся большой степенью новизны, в качестве, как правило, берётся усредненное значение.

Для радиальных сетей в задачах регрессии этого можно достичь с помощью масштабирующей функции среднее/стандартное отклонение. Обучающие данные масштабируются таким образом, чтобы среднее выходное значение равнялось 0.0, а все другие значения были бы промасштабированы на стандартное отклонение выходных сигналов. При обработке входных точек, лежащих вне областей действия радиальных элементов, выходной сигнал сети будет приблизительно равен среднему значению.

Качество работы сети в задаче регрессии можно проверить несколькими способами. Во-первых, сети можно сообщить выходное значение, соответствующее любому наблюдению (или какому-то новому наблюдению, которое необходимо проверить). Если это наблюдение содержалось в исходных данных, то выдается значение разности (невязки).

Во-вторых, могут быть получены итоговые статистики. К ним относятся среднее значение и стандартное отклонение, вычисленные для обучающих данных и для ошибки прогноза. В общем случае среднее значение ошибки прогноза будет очень близко к нулю (в конце концов, нулевое среднее для ошибки прогноза можно получить, попросту оценив среднее значение обучающих данных и вовсе не обращаясь к значениям входных переменных). Наиболее важным показателем является стандартное отклонение ошибки прогноза. Если оно не окажется существенно меньше стандартного отклонения обучающих данных, это будет означать, что сеть работает не лучше, чем простая оценка по среднему. Далее, в пакете STATISTICA Neural Networks пользователю выдается отношение стандартного отклонения ошибки прогноза к стандартному отклонению обучающих данных. Если оно существенно меньше единицы (например, ниже 0.1), то это говорит о хорошем качестве регрессии. Это регрессионное отношение (точнее, величину единица минус это отношение) иногда называют долей объясненной дисперсии модели.

В-третьих, можно вывести изображение поверхности отклика. На самом деле, разумеется, эта поверхность представляет собой $N+1$-мерный объект, где N - число входных элементов, а оставшееся измерение соответствует высоте точки на поверхности. Понятно, что непосредственно визуально представить такую поверхность при N большем двух невозможно (реально N всегда больше двух). Тем не менее, в пакете STATISTICA Neural Networks можно выводить срезы поверхности отклика по любым двум входным переменным. При этом значения всех остальных входных переменных фиксируются и меняются

только два выбранные. Всем остальным переменным можно придать любое значение по своему усмотрению (по умолчанию система STATISTICA Neural Networks возьмет для них средние значения). Значения двух исследуемых переменных можно менять в произвольном диапазоне (по умолчанию – в диапазоне изменения обучающих данных).

2.8. Аппроксимация функций в пакете STATISTICA

Рассмотрим один из методов аппроксимации функций в пакете STATISTICA. Пусть имеется неизвестная функция двух переменных, которая задана 1000 значений на интервале $[0,10]\times[0,10]$. Предполагается, что точки, в которых заданы значения функции, распределены в указанном квадрате случайно. Требуется построить математическую модель, которая описывает связь значений функции и значений независимых переменных. График функции показан на рис.2.13.

Структура данных. Файл данных состоит из трех переменных: x, y, z. Переменные x и y – независимые, переменная z – зависимая (рис.2.14).

Построение модели. В классическом подходе к решению задачи на первом шаге строится модель множественной регрессии. Очевидно, что зависимость нелинейная и все линейные модели будут давать "плохие" результаты. Необходимо построить нейросетевую модель.

Шаг 1. Запускаем нейронные сети (рис.2.15). На стартовом окне выбираем тип задачи «Регрессия». Далее переходим к выбору непрерывных входных и выходных переменных. Диалог выбора переменных с установками показан на рис.2.16. В качестве инструмента выбираем «Мастер решений» (это установка по умолчанию). Стартовое окно с описанными установками показано на рис.2.17. Нажимаем OK.

Шаг 2. На вкладке «Тип сети» (рис. 2.18) выбираем пункт «Радиальная базисная функция» (РБФ). В «Длительность анализа» установим N сетей=100 (вкладка «Быстрый», рис.2.19). Нажимаем OK.

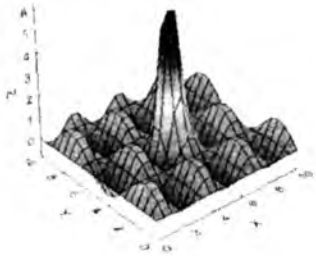


Рис.2.13. График функции

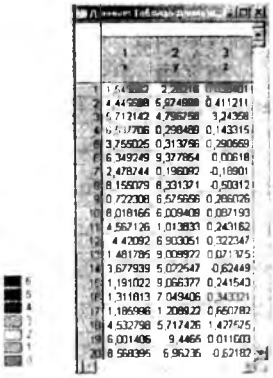


Рис.2.14. Фрагмент исходной таблицы данных

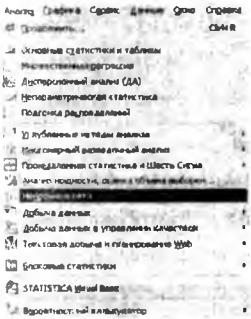


Рис.2.15. Запуск модуля «Нейронные сети»

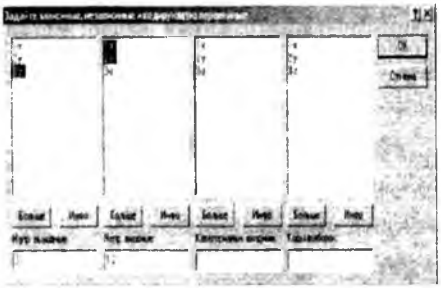


Рис.2.16. Диалог выбора переменных

Рис.2.17. Стартовое окно с проделанными установками

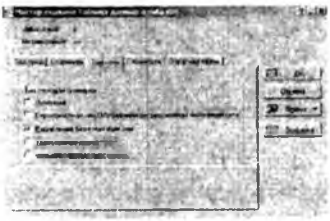


Рис.2.18. Окно задания установок «мастера решений», вкладка «Тип сети»

Шаг 3. Проанализируем результаты (рис.2.20). По мере увеличения числа элементов на скрытом слое сети увеличивается точность результатов. Ни для одной модели не получен эффект "зазубривания данных" (т.е. все модели обладают способностью обобщать результат на новые наблюдения). С помощью опции «Выбор модели» исключим все, кроме пятой сети. Построим график наблюдаемых и предсказанных значений. Для этого переходим на вкладку «Графики» и по оси X «Наблюдаемые», а по оси Y – «Предсказанные» (рис. 2.21). Строим зависимость Y от X. График показан на рис.2.22 и практически идеально лежит на прямой линии. Этого следовало ожидать, поскольку ошибки на выборках составляют всего 0.7%.

В заключение приведем граф построенной нейронной сети архитектуры РБФ (рис.2.23). Сеть имеет структуру прямого распространения с двумя входными переменными, 540 элементами на скрытом слое и одним элементом на выходном слое. Точность построенной модели равна 0.7%.

Можно сделать вывод, что сеть успешно распознала структуру обучающего множества и пригодна для использования в прогнозировании значений зависимой переменной.

Рис. 2.19. Окно задания установок «мастера решений», вкладка «Быстрый»

Рис.2.20. Окно результатов анализа, вкладка «Дополнительно»

Рис.2.21. Окно результатов анализа, вкладка «Графики»

Рис.2.22. Графики «предсказанных» и «наблюдаемых» значений

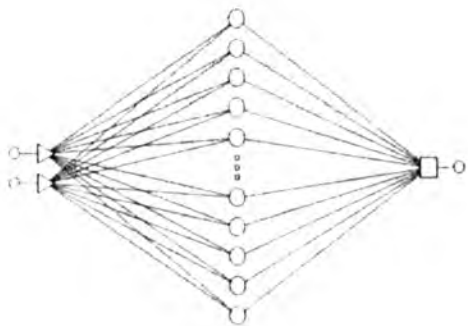


Рис.2.23. Схема построенной сети

Шаг 4. Сохранение модели. Сохранение сети производится в формате *.snn.

2.9. Модели сердечно-сосудистых заболеваний в пакете STATISTICA

Ишемия является одной из центральных проблем клинической медицины, поскольку занимает третье место среди причин смертности от неинфекционной патологии [59]. В настоящее время выделяют несколько

типов ишемии, причем тактика ведения больного в этих случаях должна быть различной. От правильности диагностики типа инсульта зависит эффективность лечебных мероприятий, а следовательно, и прогноз для каждого больного.

Для индивидуального диагноза применяют обработку больших массивов клинических данных на ЭВМ с помощью различных алгоритмов, а также привлекают инструментально-лабораторные методы. Основная цель прогноза течения и исходов сердечной патологии – это активное динамическое управление лечебно-диагностическим процессом и оптимизация его результатов. Точность прогноза существенно повышается при введении в рассмотрение большого количества информативных параметров.

В современных условиях актуальной является разработка новых диагностических критериев, математических систем прогнозирования тяжести течения и исхода болезней сердечной недостаточности. Это связано с всё более широким использованием информационных технологий в медицинской практике, усложнением и компьютеризацией медицинской диагностической аппаратуры, оптимизацией потоков информации между лабораторным отделением, вычислительным центром и палатой интенсивной терапии или операционной.

Авторами монографии предпринята попытка построения регрессионной модели разновидностей ишемической болезни сердца и математического прогнозирования на её основе с использованием нейронных сетей в пакете «STATISTICA Neural Networks 4.0» [60].

Исходная выборка была разделена на обучающую (n=268) и контрольную (n=30) с использованием рандомизации с помощью генератора случайных чисел. Предварительно с помощью генетического алгоритма нами была снижена размерность признакового пространства. В итоге в качестве исходных признаков использовалось 8 количественных показателей данных клинических и коронарных исследований (рис. 2.24). В качестве выходных 5 признаков был проставлен дифференциальный диагноз ишемической болезни сердца: норма, невыраженная ишемия, средненная ишемия, выраженная ишемия и инфаркт миокарда. Тип ишемии

(заключительный диагноз) во всех случаях верифицирован результатами комплексного клинико-инструментально-лабораторного исследования.

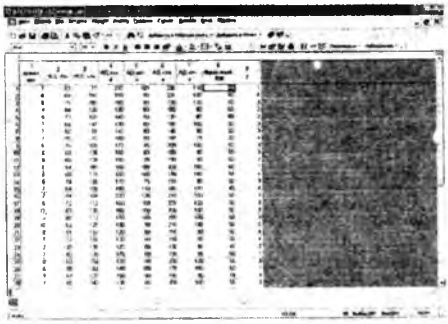


Рис.2.24. Исходные клинико-лабораторные данные дифференциального диагноза

Программный метод реализации постановки дифференциального диагноза ишемической болезни сердца по нелинейной модели регрессии осуществлён авторами монографии в пакете «STATISTICA Neural Networks 4.0» [61]. Нейронная сеть строилась и обучалась с помощью метода обратного распространения. Непрерывные выходящие, в нашем случае, – это переменная *y*. Непрерывные входящие – это переменные *x* - Время кон., ЧСС пок., ЧСС кон., АД пок. в., АД пок. н., АД кон. в., АД кон. н., Фрак. выб. ЛЖ (рис. 2.25).

Рис.2.25. Выбор переменных для «анализа»

В диалоговом окне настройки процедуры «Мастер решений» задаём алгоритм поиска нейронной сети. Имеется две возможности:

1) задать количество сетей, которые необходимо протестировать (выяснить подходят ли они для решения сформулированной задачи);

2) вручную задать время выполнения алгоритма.

Во время работы алгоритма поиска подходящей нейронной сети в этом диалоговом окне приводится информация о времени исполнения алгоритма, а также о рассмотренных нейронных сетях. Цель алгоритма поиска - перебор ряда нейросетевых конфигураций и выбор наилучшей сети.

Для того чтобы посмотреть иллюстрацию выбранной сети, нужно воспользоваться кнопкой «Архитектура сети» вкладки «Дополнительно» диалогового окна результатов.

В итоге нами была отобрана сеть - трехслойный персептрон с 8-ю нейронами на первом скрытом слое и 9-ю на втором скрытом слое, производительностью 0,257 (рис.2.26).

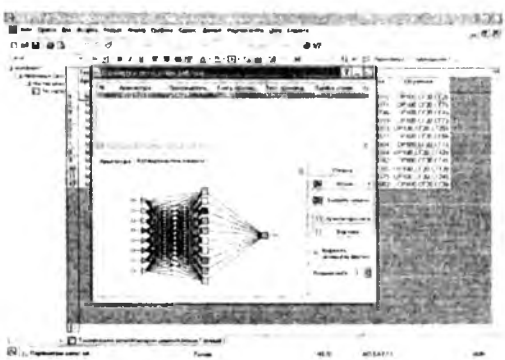


Рис. 2.26. Иллюстрация выбранной нейронной сети

Для обучения алгоритм поиска сети разбивает (по умолчанию) всё множество наблюдений на обучающее, контрольное и тестовое множества. Каждое из этих множеств несет свою важную функцию. На обучающем множестве происходит непосредственное обучение сети, т.е. изменение

весовых коэффициентов каждого из нейронов пропорционально ошибке на выходе. Соответственно все наблюдения из этого множества многократно участвуют в процедуре изменения весовых коэффициентов обучаемой сети. Наблюдения контрольного множества в процедуре изменения весов нейронов не участвуют. Основная функция этих данных в постоянном контроле способности сети к обобщению данных, на которых она не обучалась. Такая процедура называется кросс-проверкой. На каждом шаге алгоритма обучения рассчитывается ошибка для всего набора наблюдений из контрольного множества и сравнивается с ошибкой на обучающем множестве. Естественно, что эти ошибки будут отличаться. Как правило, ошибка на контрольном множестве превышает ошибку на обучающем множестве.

После того, как сеть обучена, нужно проверить качество ее работы и определить характеристики. Для этого в пакете STATISTICA Neural Networks имеется набор экранных статистик и графических средств. В том случае, если заданы несколько моделей (сетей и ансамблей), то (если это возможно) STATISTICA Neural Networks отобразит сравнительные результаты (например, построит кривые отклика нескольких моделей на одном графике, или представит предикторы нескольких моделей в одной таблице). Это свойство очень полезно для сравнения различных моделей, обучаемых на одном наборе данных.

Все статистики вычисляются раздельно для обучающего (рис. 2.27), контрольного (рис.2.28) и тестового (рис. 2.29) множеств.

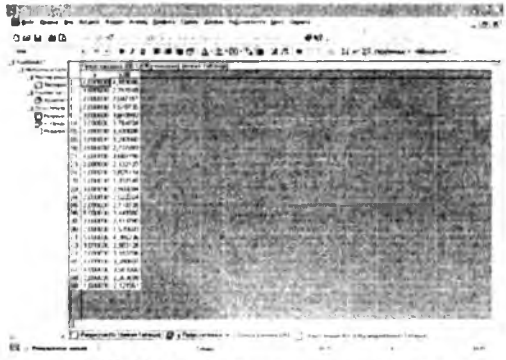


Рис. 2.27. Интерфейс обучающей выборки

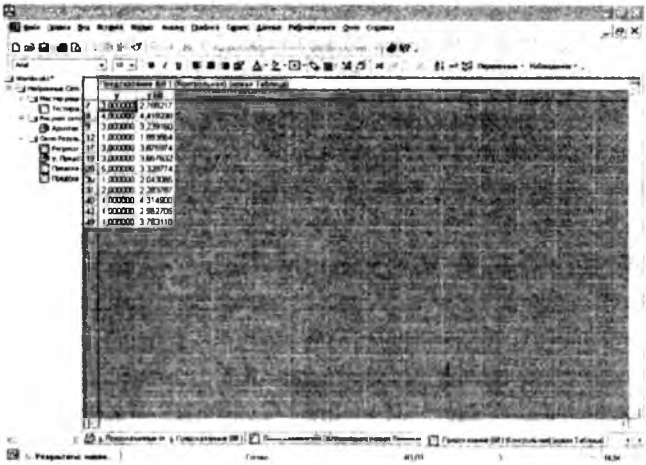


Рис. 2.28. Интерфейс контрольной выборки

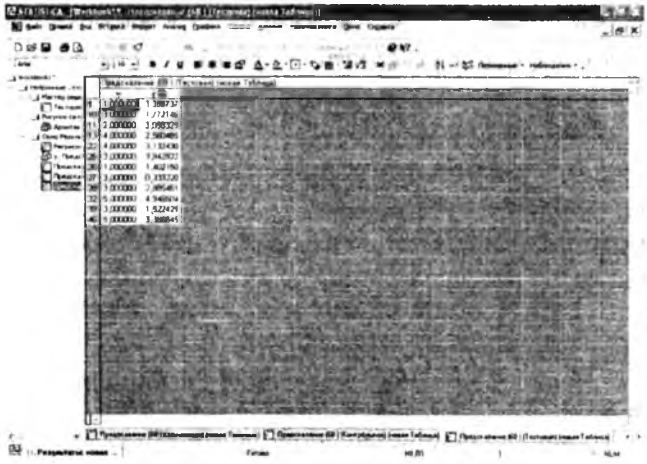


Рис. 2.29. Интерфейс тестовой выборки

Все веса и параметры активации доступны в виде удобного текстового файла, который одним щелчком мыши можно преобразовать в таблицу результатов системы STATISTICA.

Таким образом, с использованием нейросетевых технологий пакета STATISTICA Neural Networks, в частности архитектуры типа многослойной персептрон и метода обучения распространения обратной ошибки, нами построена сеть по дифференциальной диагностики ишемической болезни сердца для поддержки принятия решений и снижения доли врачебных ошибок. Характеристики полученной сети определялись с помощью статистик и графических средств. Проверка качества работы сети показала правильность постановки диагноза в 75 % случаев, что является удовлетворительным. Ошибочный диагноз при работе сети объясняется зашумлённостью исходных данных.

ГЛАВА 3. МЕТОДЫ ОБРАБОТКИ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ И РАСПОЗНАВАНИЯ ОБРАЗОВ В СИСТЕМЕ MATLAB

На сегодняшний день обработка изображений является важным направлением применения современной компьютерной технологии. Известны такие задачи обработки изображений, как фильтрация и восстановление изображений, сегментация изображений, как средства сжатия информации. Проблемы распознавания изображений кроме классической задачи распознавания фигур заданной формы на изображении ставят новые задачи распознавания линий и углов на нём, а также распознавания края изображения.

В последние годы одним из направлений развития компьютерных технологий в медицине является обработка цифровых изображений: улучшение качества изображения, восстановление поврежденных изображений, распознавание его отдельных элементов. Распознавание патологических процессов является одной из наиболее важных задач обработки медицинских изображений.

В настоящее время для решения данных задач используются компьютерные системы диагностики – CAD (computer added diagnostic). Алгоритмы CAD-системы медицинских изображений, как правило, включают в себя сегментацию изображения, выделение объектов интереса («масс»), их анализ, параметрическое описание и классификацию. При этом выделенные объекты являются отражением патологического процесса в организме, а их классификация отвечает на вопрос – все ли выделенные объекты являются проявлением патологического процесса.

В качестве методов параметрического описания выделенных объектов используется текстурный, гистограммный и морфометрический анализ. Классификация объектов интереса может проводиться по методу ронных сетей, опорных векторов, дискриминантного анализа и др. Вместе с тем, задача автоматизированной диагностики патологических

процессов по данным медицинских изображений далека от своего разрешения.

Актуальной проблемой остается определение оптимальных методов параметрического описания объектов интереса, что может оказывать непосредственное влияние на качество классификации данных объектов. Её успешное решение можно осуществить с помощью CAD-системы, в функциональную схему которой включается перевод изображения из формата DICOM в среду MATLAB с последующей автоматической сегментацией изображения различными методами и сохранением его в архиве.

3.1. Среда системы MATLAB

Появившаяся в начале 60-х годов и быстро увеличивающая свою мощность компьютерная техника позволила решать многие научные и инженерные задачи, которые не поддавались ранее решению аналитическими методами. К середине 70-х годов было накоплено большое количество алгоритмов численного решения таких задач. Образовались целые библиотеки алгоритмов. К числу лучших из таких библиотек следует отнести библиотеки матричных алгоритмов LINPACK и EISPACK департамента энергетики США. Эти библиотеки - собрание тщательно проверенных и оптимизированных алгоритмов, разработанных в течение многих лет ведущими мировыми специалистами. Они фактически представляют собой современное состояние численных методов для научных и инженерных целей. С появлением этих библиотек появилась также необходимость обеспечить кратчайший доступ широкой аудитории инженерных и научных работников к этому изобилию алгоритмов. Собственно говоря, создание языка MATLAB (сокращенно от MATrix LABoratory - матричная лаборатория) и было обусловлено этой необходимостью [62].

Когда в 1978г. Cleve Moler, разработчик системы MATLAB, предложил в качестве основного объекта языка MATLAB использовать двумерный массив (матрицу), не требующий задания размерности, возможно, он не предполагал, насколько эффективным окажется новый язык для написания матричных алгоритмов.

MATLAB за прошедшие десятилетия приобрел большую популярность, постепенно переместившись с больших вычислительных систем на персональные компьютеры, а сама программа вместе со всеми профессиональными приложениями превратилась в мощную систему, охватывающую широкий спектр научных, инженерных и экономических применений [63].

MATLAB содержит инструменты для: сбора данных; анализа и обработки данных; визуализации и цифровой обработки сигналов и изображений; создания алгоритмов и проектирования; моделирования и имитации; программирования и разработки приложений.

MATLAB выполняет множество компьютерных задач для поддержки научных и инженерных работ, начиная от сбора и анализа данных до разработки приложений. Среда MATLAB объединяет математические вычисления, визуализацию и мощный технический язык. Встроенные интерфейсы позволяют получить быстрый доступ и извлекать данные из внешних устройств, файлов, внешних баз данных и программ. Кроме того, MATLAB позволяет интегрировать внешние процедуры, написанные на языках Си, Си++, Фортран, и Java с MATLAB – приложениями.

Используемый более чем полумиллионом пользователей в промышленности, государственных, академических и учебных организациях, MATLAB фактически стал принятым во всем мире стандартом для технических вычислений. MATLAB имеет широкий спектр применений, включая цифровую обработку сигналов и изображений, проектирование систем управления, естественные науки, финансы и экономику, а также приборостроение. Открытая архитектура позволяет легко использовать MATLAB и сопутствующие продукты для

исследования данных и быстрого создания конкурентоспособных пользовательских инструментов [64].

Основные функции: быстрые и точные численные алгоритмы; графика для анализа и отображения данных; интерактивный язык и среда программирования; инструменты для настройки пользовательских интерфейсов; интерфейсы с внешними языками, такими как Си, Си++, Фортран, и Java; поддержка импорта данных из файлов и внешних устройств плюс доступ к базам данных и вспомогательному оборудованию при помощи приложений; преобразование MATLAB-приложений в Си и Си++ при помощи набора Compiler Suite. Этот широкий набор возможностей делает MATLAB идеальной базой для решения технических проблем (рис.3.1).

Среда MATLAB спроектирована для интерактивных или автоматических вычислений. Использование встроенных математических и графических функций и простых в использовании инструментов позволяет легко анализировать и отображать данные. Структурированный язык и программные инструменты позволяют сохранять результат интерактивных исследований, разрабатывать собственные алгоритмы и приложения.

Надежные математические функции и численные расчеты. С более чем шестью сотнями математических, статистических и инженерных функций MATLAB предоставляет немедленный доступ к высокопроизводительным численным расчетам. Численные алгоритмы быстрые, точные и надежные. Эти алгоритмы, разработанные экспертами в математике, являются фундаментом языка MATLAB. Ядро математического процессора включает библиотеки алгоритмов линейной алгебры LAPACK и BLAS, а также библиотеку цифровой обработки сигналов FFTW, внедряющую последние достижения в математических расчетах непосредственно в MATLAB.

Математика оптимизирована для матричных и векторных операций, что позволяет использовать ее вместо языков более низкого уровня,

подобных Си и Си++, получая ту же скорость вычислений при значительной экономии

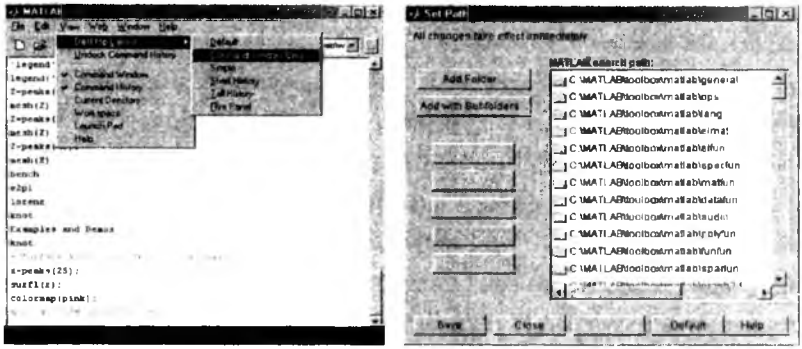


Рис.3.1. MATLAB- интерфейс – быстрый доступ к функциям MATLAB и сопутствующим продуктам

времени на программирование. Благодаря обширному набору оптимизированных математических алгоритмов, встроенных в MATLAB, разработчики могут сконцентрироваться на принципиальных аспектах проекта, не теряя времени на поиск, разработку и отладку известных алгоритмов.

MATLAB содержит много функций для осуществления математических вычислений и анализа данных, включая:

- Линейную алгебру и арифметику матриц, линейные уравнения, собственные значения, сингулярные значения и матричные разложения; полиномы и интерполяцию: стандартные полиномиальные операции, такие как вычисление корней полиномов, дифференцирование, аппроксимация кривыми и разложение функций; обработку сигналов: цифровые фильтры, быстрое Фурье преобразование (FFTs) и свертка; анализ данных и статистику: описательная статистика, предварительная обработка данных, регрессия, аппроксимация кривыми, фильтрация данных.

- MATLAB-функции, работающие не с массивами данных, а с математическими функциями, включая функции построения графиков, оптимизации, нахождения нулей и численного интегрирования;

- Дифференциальные уравнения: решение дифференциальных уравнений, включая задачи с начальными значениями для обыкновенных дифференциальных уравнений и дифференциально-алгебраические уравнения, а также задачи с граничными условиями и смешанного типа для систем эллиптических и параболических уравнений с частными производными;

- Разреженные матрицы – включая как специальные так и общие математические операции, в том числе итеративные методы для разреженных систем линейных уравнений.

Дополнительные пакеты-расширения (тулбоксы – toolboxes) значительно увеличивают мощность MATLAB;

Тулбоксы – наборы MATLAB-функций и интерфейсов - позволяют использовать значительно больше математики и вычислительной мощности при решении различных задач. Распространяемые тулбоксы включают приложения для цифровой обработки сигналов, оптимизации, статистики, символьной математики, решения уравнений в частных производных, аппроксимации данных и многие другие [64].

Качественная графика помогает проиллюстрировать и лучше понять результаты численных расчетов.

MATLAB обеспечивает немедленный доступ к специализированным графическим функциям, включая:

- Двух- и трехмерные графики, такие как линейные, логарифмические, гистограммы, графики поверхностей различного типа (проволочный каркас, гладкие, кусочно-гладкие и др.);

- Поддержка триангулированных и решетчатых данных;

- Объемная визуализация для наблюдения скалярных и векторных данных;

- Вывод изображений на экран и в файл;

- Интерактивное редактирование и аннотирование графиков;

- Аппаратная и программная поддержка OpenGL;

- Различного вида диаграммы: рассеяния, столбчатые, секторные и др.;

- Анимация (мультипликация) и звук;

- Различные источники освещения для цветных поверхностей;

- Наблюдение при помощи камер и управление перспективой;

- Интерактивное и программируемое управление индивидуальными атрибутами графиков, такими как линии, оси, легенды и параметры страницы;

- Освещение прямым источником, по методу Гуро и Фонга.

Графическая часть MATLAB представлена специализированными графиками, используемыми в науке и технике, помогающими понять сложные системы, начиная от линий на плоскости и контурных графиков и кончая интерактивными графическими пользовательскими интерфейсами (рис.3.2). Сюда входят:

- Инструменты построения при помощи мыши графического пользовательского интерфейса и поддержки API;

- Импортирование графических файлов известных форматов, таких как EPS, TIFF, JPEG, PNG, BMP, HDF, AVI, и PCX;

- Печать и экспорт графики в другие приложения, такие как Word и PowerPoint, в различных популярных форматах;

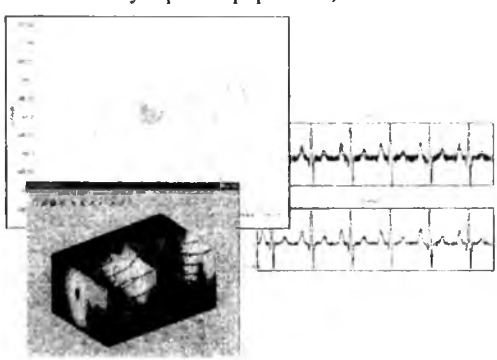


Рис.3.2. Специализированные графики MATLAB

- Расширенная поддержка цифровой обработки изображений и приложения географического картографирования при помощи тулбоксов.

MATLAB содержит полный набор инструментов для визуализации трехмерных скалярных и векторных данных, включая функции создания изоповерхностей и линий тока. Эти функции предназначены, чтобы помочь ученым и инженерам наблюдать и понимать большие, часто сложные многомерные данные.

Дополнительные приложения, такие как Image Processing Toolbox и Mapping Toolbox, расширяют возможности MATLAB для применения, включая цифровую обработку данных и географическую картографию.

Унифицированный интерактивный язык и среда программирования. Язык MATLAB разработан для интерактивных и автоматических вычислений. С оптимизированными матричными функциями разработчик может осуществить интерактивный анализ, в то же время языковые возможности структурирования программ позволят ему эффективно разрабатывать собственные алгоритмы и приложения. Универсальный язык позволяет справиться с множеством задач, включая сбор данных, анализ, разработку алгоритмов, имитацию систем и разработку приложений.

Возможности языка включают структуры данных, объектно-ориентированное программирование, инструменты разработки графического пользовательского интерфейса, функции отладки и возможность подключать ваши C, C++, Fortran, и Java процедуры.

Упрощение кода за счет векторизации. Многие математические функции MATLAB оперируют непосредственно с векторами и матрицами. В результате задачи, которые могли бы включать десятки строк и требовали интенсивных вычислений с циклами на языке Си, требуют вызова только одной функции в MATLAB. Более короткие программы требуют меньше времени на их выполнение и легче настраиваются.

Удобный язык программирования: пользователь использует тот же самый язык в MATLAB при интерактивных вычислениях и при разработке

программ в редакторе-отладчике. Эта совместимость позволяет изучать только один язык и синтаксис, интерактивная работа может быть сохранена и использована для будущих проектов без повторного программирования.

Инструменты, подобные редактору-отладчику М-файлов, а также проектирования графического пользовательского интерфейса позволяют разрабатывать приложения и настраивать графику. Разработчики MathWorks, используя инструменты разработки графического пользовательского интерфейса MATLAB, создали SPTool, визуальный интерфейс, который поставляется вместе с Signal Processing Toolbox (рис.3.3).

Инструменты проектирования. MATLAB включает множество инструментов для поддержки разработки аналитических процедур, алгоритмов и законченных приложений. Эти инструменты включают:

- Командное окно для интерактивных вычислений и проектирования;

- Окно Command History для записи всех команд в течение интерактивной MATLAB сессии;

- Инструмент Current Directory для нахождения и навигации по файлам и директориям.

- Редактор - отладчик М-файлов, который позволяет:

- разрабатывать и отлаживать М-файлы (функций и сценариев);

- выделять различными цветами резервные слова и комментарии;

- выполнять пошаговое выполнение кода;

- проверять скобки и нумерацию строк.

- осуществлять установку контрольных точек, остановку на ошибке или на предупреждении;

- вычислять переменные по месту нахождения.

Кроме этого, среда предоставляет инструменты для интерактивного создания и редактирования пользовательского интерфейса, включающие окна списков, ниспадающие меню, кнопки и переключатели, полосы прокрутки и другие стандартные элементы интерфейса.

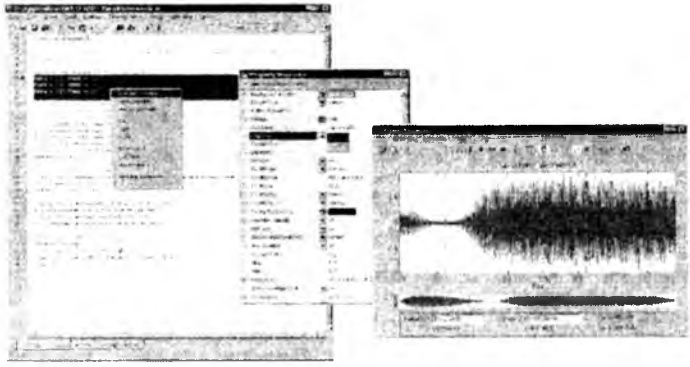


Рис.3.3. Инструменты для создания приложений в MATLAB

Имеются следующие инструменты:

- Layout Editor – добавление и настройка объектов в окне;

- Alignment Tool – выравнивание объектов по отношению друг к другу;

- Property Inspector – проверка и установка значений для свойств объектов;

- Object Browser – наблюдение иерархического списка объектов Handle Graphics в текущей MATLAB сессии.

Таким образом, подытоживая вышеописанное, можно утверждать, что популярная и мощная система компьютерной математики MATLAB отражает современный уровень моделирования, обработки и фильтрации разнообразных по своей природе данных.

3.2. Обработка биомедицинских изображений с использованием системы MATLAB

Анализ рентгенограмм, очень распространенная диагностическая операция, характеризуется высоким общим процентом ошибочных диагнозов. Это обусловлено тем, что рентгенорадиологические изображения - теневые двухмерные изображения трехмерных тел. Их

специфическим свойством является взаимное наложение теней объектов, расположенных по ходу излучения, что затрудняет интерпретацию изображения. Поэтому при исследовании конкретный орган стремятся контрастировать, вводя в него препараты, резко меняющие относительно окружающих объектов его радиационную прозрачность. Вместе с тем ясно, что необходимая информация содержится и в неконтрастированном изображении и можно попытаться сделать её доступной специалисту, применяя соответствующие методы обработки.

Довольно часто рентгеновские биомедицинские изображения не отвечают тем критериям качества, которые необходимы для их достоверного анализа. Также не всегда существует возможность сделать повторный снимок. Это приводит к необходимости цифровой обработки такой информации [65]. Основными недостатками рентгеновских изображений в большинстве случаев являются искаженные яркостные характеристики и низкая контрастность. Рассмотрим пример обработки одного из таких изображений с помощью системы MATLAB.

Недостаток исходного биомедицинского изображения (рис.3.5 а) состоит в том, что это изображение низкоконтрастное, что затрудняет анализ мелких деталей. Поэтому сначала выполняется операция растяжения гистограммы изображения на максимально допустимый диапазон (рис.3.5 б). Далее осуществляется контрастирование исследуемого изображения (рис.3.5 в). Это приводит к улучшению визуального качества рентгеновских изображений.

а) б) в)

Рис. 3.5. Обработка рентгеновских изображений в среде Matlab

На практике, конечно, применяются также и другие более сложные методы и алгоритмы обработки изображений такого рода.

Приведём пример программы обработки биомедицинских изображений в среде MATLAB

```
L=imread('cardial.bmp');
figure, imshow(L);L1=imadjust(L,[min(min(L)) max(max(L))]/255,[],1);
figure, imshow(L1);L=L1(:,:,1);
L=double(L);
Filter=[1 1 1,1 1 1,1 1 1]; Lser=filter2(Filter,L(:,:,1),'same')./9;
C=abs(L(:,:,1)-Lser)./(L(:,:,1)+Lser);
C=C.^.55;
[N M]=size(L);for i=1:N;disp(i);
for j=1:M;if L(i,j,1)>Lser(i,j);
Lvyh(i,j)=Lser(i,j)*(1+C(i,j))/(1-C(i,j));
else
Lvyh(i,j)=Lser(i,j)*(1-C(i,j))/(1+C(i,j));
end;
end;
end;
figure, imshow(Lvyh/255);
```

3.3. Синтез электрокардиосигналов в системе MATLAB

Сфера высоких технологий, которая сегодня все глубже проникает в области исследования различных биологических процессов и диагностики заболеваний, предоставляет новые возможности для исследователей. Появление современной электрорегистрирующей аппаратуры, новых биофизических и математических методов исследования предоставляет эффективные диагностические способы для регистрации и анализа электрической активности сердца.

Важность, значимость и актуальность клинической электрокардиографии очевидны, поскольку исследования биоэлектрической активности сердца являются сегодня незаменимыми в диагностике нарушений ритма и проводимости, гипертрофий желудочков и предсердий, ишемической болезни сердца, инфаркта миокарда и других кардиозаболеваний. Анализ биоэлектрических процессов, протекающих в сердце, ведется главным образом при помощи электрокардиограммы, которая представляет собой разность биоэлектрических потенциалов работающего сердца. Именно форма и количественный состав ее компонент позволяют диагностировать норму или патологию в сердечной деятельности [66].

Одним из направлений математического моделирования в области электрокардиографии является создание самой электрокардиограммы. Этот подход можно считать удачным и верным, потому что проведение натурных экспериментов может быть затруднено по ряду объективных причин: несовершенство регистрирующего оборудования или его дефицит, частое отсутствие возможности длительного мониторного слежения. Полученные результаты при подобном подходе можно считать адекватными реальным данным с некоторой долей погрешности и применять на практике [67].

Исследование ЭКГ-сигналов трех стандартных отведений от конечностей очень часто применяется при диагностировании. Поэтому особое внимание стоит уделить моделированию именно такого рода сигналов. Полученные модели могут найти в практике широкое применение, начиная от калибровки электрокардиографов и заканчивая исследованиями моделей на предмет выявления особенностей сигналов в определенных клинических картинах.

Авторы монографии для моделирования электрокардиографических сигналов (ЭКС) [68] с использованием стандартных отведений применили метод линейно-квадратичной интерполяции [69]. В рамках этого метода ЭКС задается в виде последовательности отрезков и участков парабол

(элементов ЭКС), которые соответствуют реальным сегментам и зубцам сигнала (рис. 3.6).

Разработанные алгоритмы синтеза электрокардиосигналов на основе линейно-квадратичной интерполяции реализованы в системе MATLAB 6.0 Release 12 фирмы The MathWorks, Inc.

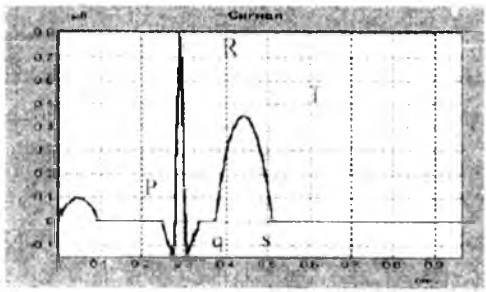


Рис. 3.6. Электрокардиограмма сердечного цикла

Алгоритм имеет MATLAB-прототип, который может быть использован независимо или в составе графического интерфейса пользователя (ГИП), реализованного в среде MATLAB (рис. 3.7). ГИП является достаточно простым с точки зрения пользования инструментом и позволяет оценивать или исследовать характеристики элементов ЭКС.

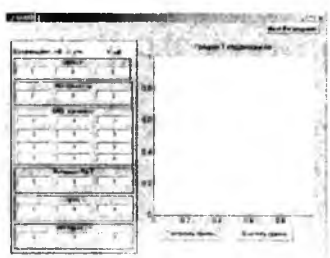


Рис.3.7. Графический интерфейс пользователя

Рис.3.8 иллюстрирует, как выглядит ГИП после окончания моделирования и нажатия кнопки вывода графика. График имеет масштаб по оси «*X*», определяемый длиной отображаемого вектора данных. Кроме того, график можно просмотреть на выбранном участке в пределах этой длины. ЭКГ может быть также отображена в автоматически выбираемом масштабе по оси «*Y*». На графике можно убирать или восстанавливать сетку. График ЭКС, воспроизводимый в правой части ГИП, можно дополнительно выводить в отдельном стандартном окне для последующего копирования, сохранения или печати.

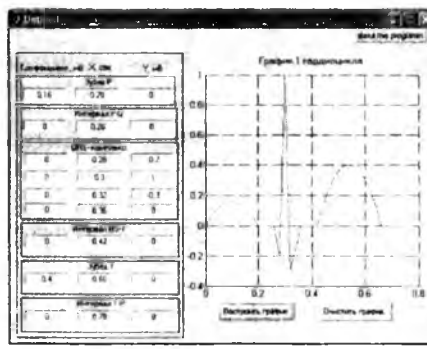


Рис.3.8. Электрокардиосигнал на основе линейно-квадратичной интерполяции

С целью описания изменения формы и очередности появления элементов ЭКС применялась система координат (СК). Момент времени отклонения первого элемента ЭКС от собственной изоэлектрической линии является началом набора СК в текущем кардиоцикле. Изоэлектрическая линия ЭКС является осью абсцисс, соответствующей этому сигналу СК. Таким способом формируется цикл ЭКС, относительно которого определяются координаты исходных и конечных точек элементов ЭКС.

Вычисление параметров элементов ЭКС способом линейно-квадратичной интерполяции включает в себя три этапа:

- задание желаемых исходных и конечных координат элементов, значений ординат экстремумов парабол и формирование последовательности триплетов;
- вычисление для каждого элемента параболы абсциссы её экстремума;
- формирование множества точек элементов ЭКС.

В синтезированном ЭКГ-сигнале сохранены основные частотные составляющие, свойственные QRS-комплексам. Программа написана в форме своеобразного электронного методического пособия. Возможности среды программирования MATLAB позволили реализовать сочетание большого количества теоретической информации и расчетных узлов программы, где каждый пользователь имеет возможность изменять значения параметров характерных элементов ЭКГ-сигнала, наблюдая за изменениями формы его сегментов. Главным достоинством алгоритма является возможность синтеза произвольных ЭКГ-сигналов, то есть вид, последовательность и форма отдельных элементов ЭКС зависят только от желания пользователя.

Авторами монографии рассматривался также метод моделирования ЭКС стандартных отведений с использованием кубических сплайнов [70], реализованный в системе MATLAB 6.0.

Система научных и инженерных расчетов MATLAB 6.0 имеет в своем составе множество пакетов прикладных программ и функций, которые обеспечивают решение различных специфических задач. В частности, для интерполяции и аппроксимации данных существуют следующие базовые программные средства:

- Spline Toolbox (содержит набор функций для работы со сплайнами);
- PolyFun Toolbox (включает в себя функции для аппроксимации и интерполяции данных).

Для реализации алгоритма использовались следующие предположения:

– координаты точек, составляющих модель ЭКГ-сигнала, определяются пользователем;

– временная составляющая сигнала (абсцисса точек) должна монотонно возрастать и не иметь разрывов.

Главным достоинством алгоритма является возможность синтеза произвольных ЭКГ-сигналов, т.е. вид, последовательность и форма отдельных элементов ЭКС зависит только от желания пользователя (он буквально «рисует» сигнал).

Ввод точек осуществлялся при помощи функции ginput(...). Его окончанием считался щелчок левой кнопкой мыши за границей графической области.

Для нахождения точек, принадлежащих сигналу, и отличающихся постоянством шага по оси времени использовалась функция INTERP1.

Рассматриваемая функция имеет следующий синтаксис:

Y_i = interp1 (x, y, X_i, '<метод>')

Эта функция строит интерполирующую кривую указанным методом для одномерного массива *y*, заданного на сетке *x*; выходной массив Y_i может определяться на более мелкой сетке X_i. Экспериментальным путем установлено, что наилучшее качество интерполяции достигается при использовании метода PCHIP. Он реализует кусочно-кубическую интерполяцию методом Хермита. Хорошее качество полученных результатов

обеспечивается за счет свойств используемого метода:

– интерполированная функция $P(x)$ представляет собой различные кубические функции на подинтервалах $x_i \le x \le x_{i+1}$;

– первая производная $P'(x)$ непрерывная;

– вторая производная $P''(x)$ кусочно-линейная.

Именно из-за этих свойств наиболее точно обеспечивается построение ЭКГ-сигнала, так как он имеет в своем составе участки изломов (например,

QRS-комплекс). Использование метода PCHIP не приводит к их сглаживанию по причине отсутствия непрерывности второй производной. Кроме этого, использованный метод не приводит к осцилляции сигнала даже при несглаженности данных.

Окно модуля и схема алгоритма синтеза ЭКС на основе табличной интерполяции и применения кубических сплайнов приведены на рис. 3.9 и рис.3.10 соответственно.

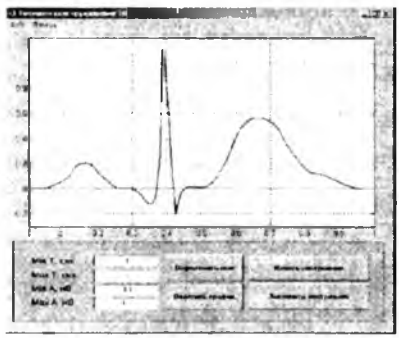


Рис.3.9. Окно модуля «Точечное конструирование ЭКС»

Рис.3.10. Схема алгоритма синтеза ЭКС
с использованием кубических сплайнов

Таким образом, построение одного цикла ЭКГ-сигнала на основе рассмотренного метода имеет следующие этапы:

– формирование точек сигнала пользователем;

– применение функции INTERP1(..., 'PCHIP') к полученным результатам для получения точек сигнала на более мелкой равномерной сетке.

Тесное сотрудничество кафедры «Приборостроение» с Республиканским специализированным центром хирургии им. академика В.Вахидова позволяет студентам и магистрантам, обучающимся на кафедре по направлению «Биомединженерия», проводить научные исследования различных биологических процессов со сходным ЭКГ-сигналу спектральным составом.

3.4. Анализ сцинтиграмм печени с использованием среды MATLAB

По данным Всемирной организации здравоохранения рак, как причина смертности людей, занимает второе место после болезней сердца. Лечение больных злокачественными опухолями является одной из наиболее актуальных и трудных проблем современной медицины, при этом успех лечения во многом определяется правильной и своевременной диагностикой. Одним из наиболее перспективных методов обнаружения злокачественных образований являются методы радиоизотопной диагностики, в частности сцинтиграфии (гамма-топография) печени.

Разработаны многочисленные методики радионуклидного исследования метаболизма печени, кровотока и внесосудистого плазмотока в ней, оценки артериопортального и портосистемного шунтирования. Созданы разнообразные гепатотропные радиофармпрепараты. В зависимости от того, какой из них применен (радиоактивный коллоид, меченый краситель, меченая аминокислота и т. ч.) можно судить о различных функциях печени (барьерной, желчевыделительюй и пр.). Современные гамма-камеры позволяют

исследовать распределение радиофармпрепарата в печени, т. е. получать изображение органа на сцинтиграмме и изучать динамику накопления и выведения препарата из любого участка печени. Патологические образования, накапливающие радиофармпрепарат сильнее, чем нормальная ткань печени, обусловливают на сцинтиграмме участки повышенной интенсивности (горячие очаги). Большинство опухолей, а также кисты и абсцессы обнаруживаются как участки пониженной интенсивности (холодные очаги). Но особенно ценны радионуклидные исследования для оценки функций печени при диффузных поражениях, для дифференцировки хронических гепатитов и циррозов, для изучения динамики функций при лечении больного. Методы радиоизотопной сцинтиграфии печени позволяют получить с минимальной опасностью для пациента значительную информацию о состоянии органа на основании изучения картин распределения радиоизотопного индикатора, представляющих собой полутоновые изображения.

Цирроз печени сопровождается серьёзными нарушениями печёночной гемодинамики и различных функций печени. Количественная оценка этих нарушений является основой установления диагноза и прогноза заболевания. В связи с этим понятен интерес к применению различных радионуклидных методов исследования функций печени и печёночной гемодинамики у больных циррозом печени [71, 72]. Однако в повседневной практике нет реальной возможности принять весь комплекс радионуклидных методик, в связи с чем важно знать о наиболее информативных радионуклидных тестах для определения состояния печени и оценки эффективности лечения. Этим целям наиболее соответствует сцинтиграфия печени с Tc-99m бутилидой. Несмотря на то, что радиоизотопные методы исследования нашли широкое применение в медицинской практике, до настоящего времени наибольшее распространение получила интерпретация сцинтиграмм печени, основанная на визуальном анализе. При этом целый ряд субъективных и объективных факторов вызывает 30 % ошибок при классификации врачом

сцинтиграмм печени класса "норма" и 6,5 % ошибок при классификации сцинтиграмм класса "патология". Наличие субъективного фактора приводит к необходимости автоматизации процесса формирования диагностических заключений.

Авторами монографии проведены анализ и обработка данных динамической гепатосцинтиграфии с Tc-99m бутилидой [73, 74] со 115 больными циррозом печени с портальной гипертензией. Среди них 72 – мужчины и 43 – женщины, в возрасте 36,4±4,2 лет. Все больные, в зависимости от тяжести цирротического процесса и степени печёночной недостаточности, были разделены на три функциональных класса, согласно классификации, предложенной Child (1964). Параллельно использовалась классификация М.Д. Пациора (1974), отражающую степень гемодинамических нарушений. В класс «А» с начальными признаками печеночной недостаточности и компенсированной портальной гипертензией вошли 19 больных. 56 человек составили класс «В» с умеренными проявлениями печёночной недостаточности и субкомпенсированной портальной гипертензией. Остальные 40 человек составили группу с тяжёлыми проявлениями печёночной недостаточности и с декомпенсированной портальной гипертензией и были отнесены к группе «С». Контрольную группу составили 16 человек.

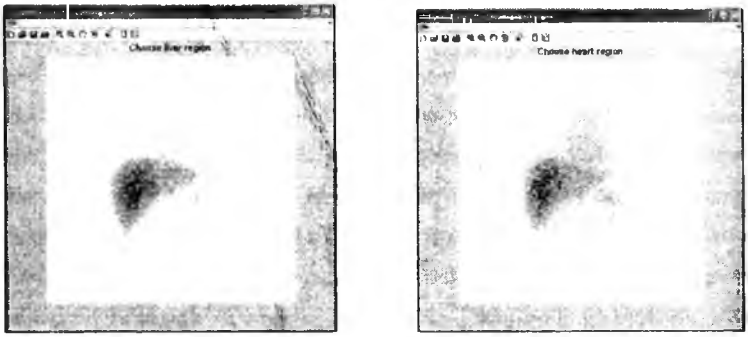


Рис. 3.11. Сцинтиграмма печени, полученная с помощью однофотонного эмиссионного компьютерного томографа

Динамическую гепатосцинтиграфию проводили на гамма-камере (FO Gamma LFOV с компьютером PDP 11/34, США). Для исследования поглотительно-экскреторной функции печени применялась методика радионуклидной динамической гепатосцинтиграфии с радиофармпрепаратом 99mTc-бутилида. При анализе сцинтиграмм (рис.3.11) широко применяются математические методы, системный анализ, камерное моделирование физиологических и патологических процессов. При анализе гепатосцинтиграмм авторы монографии отошли от стандартной методики их обработки, применив методику P.H. Brown (1988), позволяющую математически разделить поглотительно-секреторную функцию печени на две составляющие, не зависящие друг от друга. Данная методика основана на том, что при обработке радиохронограмм применяется деконволюционный анализ с использованием прямого и обратного преобразования Фурье.

При использовании преобразований Фурье изображение представляется в виде суммы сложных показательных функций переменных амплитуды, частоты и фазы. Преобразование Фурье играет очень важную роль во многих областях обработки изображений, включая улучшение, анализ, восстановление и сжатие. Основной смысл деконволюционого анализа сводится к получению количественной оценки печёночной фракции экстракции (Fr extr, %), отражающей поглотительную способность гепатоцитов. При этом применяется кривая время-активность (T/A) правой доли печени как выход и кривая T/A сердца как вход. Кривая T/A печени была сглажена перед деконволюцией функцией $y(t) = at^b \cdot e^{ct}$ ввиду наличия случайного шума. Кривая истинного отклика печени была деконволирована из входной и выходной кривых, применяя метод преобразования Фурье с длинным хвостом, прибавленным к концу входной и выходной кривых. Формула для прибавления хвоста в кадре t кривой сердца и печени выглядит следующим образом:

$$TAiL = 0{,}5 \cdot AMP \cdot (\cos[\pi \ (t \ \ 30)/(128 \ \ 30)] + 1.$$

где i=30-128; AMP – величина амплитуды на 30-ом кадре.

Далее выполняется прямое и обратное преобразование Фурье для получения истинного отклика печени. Гепатоцитная фракция экстракции определялась путём экспоненциальной подгонки, которая заключалась в том, что проводилась экстраполяция экспонентой методом наименьших квадратов в порядке обратно-пропорциональном времени от 30-ого кадра назад к первым кадрам и выражалась соотношением:

$$HEF = \frac{Y_{exp}}{Y_{max}} \cdot 100\%,$$

где Y_{exp} - максимальная величина экспоненциально-найденной кривой; Y_{max} - максимальная величина деконволюционной кривой.

Выделительная функция определялась по оценке кривых время-активность (T/A), которые были получены из правой верхней доли печени и селезёнки. Кривая T/A селезёнки была отнесена к фоновой и вычтена из кривой T/A печени.

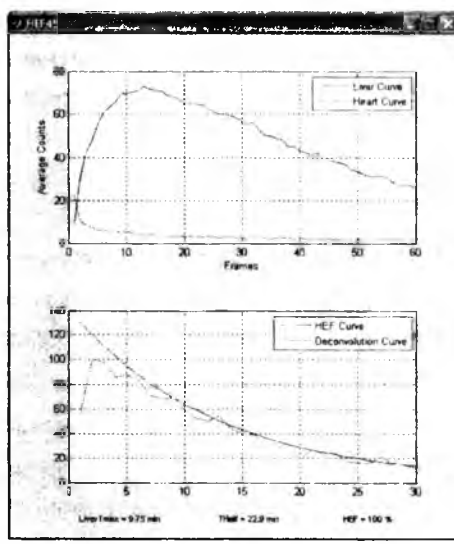


Рис. 3.12. Результаты, полученные в MATLAB

Таким образом, при использовании деконволюционного анализа при обработке радиохронограмм были получены два показателя, характеризующие функциональную активность печёночной паренхимы: Frextr – показатель фракции экстракции (%), характеризующий поглотительную активность гепатоцитов; T1/2 – период полувыведения радиофармпрепарата (мин), характеризующий выделительную функцию печени (рис.3.12).

Количественный анализ кривых подразумевает использование различных параметров, которые характеризуют: 1) высоту ординаты кривых (т.е. их амплитуду); 2) временные интервалы между отдельными («характерными») точками кривых, например, началом исследования и максимальной высотой подъема, половиной спада кривой и т. д.; 3) величину площади под кривой или какой-то ее частью (например, на этапе подъема или, наоборот, спуска).

При оценке радиографических кривых на основе количественного анализа надо помнить, что все параметры кривых (временные, линейные, характеризующие площадь, абсолютные и относительные) являются характеристиками самих кривых, а не собственно функций органа. Используя компьютерную обработку кривых, врач обязан вдумчиво и корректно оценить состояние органа, сопоставляя радионуклидные данные с результатами биохимических, радиохимических и других исследований и всегда имея в виду клиническую картину заболевания.

Наиболее точным и физиологически обоснованным способом анализа радиографических кривых является их интерпретация на основе математического моделирования транспорта РФП в соответствующих системах и органах. Для этой цели нами созданы специальные пакеты прикладных программ, которыми должны оснащаться лаборатории радионуклидной диагностики.

Разработанное нами программное обеспечение предназначено для сцинтиграфии печени с Tc-99*м* мезидой. Его можно применять в динамической гепатосцинтиграфии с деконволюционным алгоритмом

обработки гепатосцинтиграмм. Программа позволяет проводить мониторинг функциональной активности печёночной паренхимы с целью выявления печёночной недостаточности как с диагностической, так и с прогностической целью, что особенно важно в хирургической клинике при лечении пациентов циррозом печени с портальной гипертензией.

Использование программного обеспечения предоставляет наиболее полную информативность в определении глубины гепатодепрессии и тяжести циррозического процесса для однофотонного эмиссионного компьютерного томографа (ОЭКТ) фирмы Medico Nuclear Spirit.

ГЛАВА 4. ЭКСПЕРТНЫЕ СИСТЕМЫ В МЕДИЦИНЕ

Развитие новых методов диагностики, а также расширение возможностей уже существующих методов было и остаётся актуальной задачей в медицине. Появление новых диагностических и лечебных технологий требует использования методов искусственного интеллекта для обработки и интерпретации данных с возможностью накопления, хранения и многократного использования медицинских данных.

Одними из наиболее эффективных средств в данной области являются экспертные системы. Они позволяют автоматизировать процесс принятия решения при освидетельствовании пациентов, повышая уровень квалификации пользователя до уровня опытных экспертов [75].

Поэтому необходимо, чтобы ЭС обладали возможностью гибкой постановки задач, были применимы ко всем областям биологии и медицины, обладали большой информационной ёмкостью и помехоустойчивостью, не нуждались в длительном времени для разработки.

Непрерывное развитие средств вычислительной техники расширяет потенциальные возможности подобных систем, в связи с чем необходимо постоянно пополнять знания в данной области. ЭС в течение длительного времени применяются в медицине для диагностики заболеваний. Каждая система при этом обладает ограниченной областью применения из-за изначальной направленности разработки.

В настоящее время медицинские диагностические системы строятся на основе нескольких методов обработки информации:

- построение дерева решений;
- статическая обработка данных;
- использование элементов искусственного интеллекта [76].

4.1. Экспертные системы – эффективное диагностическое средство в медицине

В медицине врачи принимают важное решение, которое определяет успех всей работы: ставят диагноз. Точность диагностики зависит от квалификации специалиста (эксперта) — его умения правильно проанализировать имеющуюся информацию. Но бывают ситуации, когда нет высококвалифицированного специалиста по какой-либо специальности. Поэтому, по мере развития вычислительной техники, возникла идея заложить знания специалистов в компьютер и использовать его в качестве электронного эксперта.

По способу решения задачи диагностики различают вероятностные системы и экспертные системы. В вероятностных системах диагностика осуществляется реализацией одного из методов распознавания образов или статистических методов принятия решений. В экспертных системах реализуется логика принятия диагностического решения опытным врачом.

Так как ЭС основана на знаниях – те, в свою очередь, должны быть явными и доступными, высококачественными, полученными непременно на уровне профессионалов, более глубокими и полными, чем у обычного пользователя, и, таким образом, должны обеспечивать глубину и компетентность системы. Система также должна уметь работать не только с количественной, но и с трудно формализуемой качественной информацией, что особенно важно для решения медицинских задач.

Разрабатываемые в настоящее время медицинские экспертные системы просты и решают узкоспециализированные задачи медицинской диагностики. По сути дела это диалоговые базы данных, сопряженные с базами знаний и подсистемами генерации отчетов.

Самые важные области применения ЭС в медицине:

- неотложные и угрожающие состояния;
- дефицит времени;
- ограниченные возможности обследования;

- скудная клиническая симптоматика;

- быстрые темпы развития заболевания.

Необходимо отметить, что работа с ЭС может вестись удаленно. Общий принцип, положенный в основу формирования ЭС диагностических заключений, – включение в базу знаний синдромов, позволяющих контролировать все основные системы организма.

Выводы, основанные на опыте работы с ЭС, весьма конкретны и обоснованы:

- структурное представление медицинских знаний в виде иерархически организованных описаний;

- этап формирования эталонных описаний развивает у молодых специалистов навыки систематизированного и логически обоснованного описания и представления клинической картины заболевания;

- язык, встроенный в систему для описания заложенных в оболочку ЭС алгоритмов вывода заключений, должен быть прост и доступен практикующему врачу.

Основные участники создания медицинских экспертных систем: врач; математик; программист.

Основная роль в разработке такой системы – роль врача – постановщика задач. Такая расстановка обязанностей не уменьшает значение разработчиков ЭС, а лишь упорядочивает взаимоотношения между участниками совместной работы.

Прежде всего, важно правильно определить, в какой форме лучше хранить знания эксперта и как ими пользоваться. Также важно обеспечить правильное применение знаний, позволяющее сформулировать достоверные выводы на основе часто противоречивой исходной информации. Желательно, чтобы система по отдельному запросу объясняла свою линию рассуждения в понятном пользователю виде. Хорошая экспертная система имеет блок для пополнения базы знаний.

Таким образом, у полностью оформленной ЭС присутствуют четыре основных блока.

- База знаний;
- Машина вывода;
- Модуль извлечения знаний;
- Система объяснения принятых решений.

Блоки 3 и 4 необязательны, но законченная ЭС обязана иметь их все.

Рассмотрим теперь основные блоки ЭС подробнее.

База знаний. База знаний содержит факты или утверждения и правила. Факты являются краткосрочной информацией, они могут изменяться в ходе одного сеанса работы. Правила составляют долговременную информацию о том, как порождать новые факты на основе известных данных. Отличие базы знаний от базы данных состоит в механизме пополнения информации недостающими фактами.

Распространенным методом отображения знаний являются правила продукций. При этом правила имеют вид ЕСЛИ – ТО, например, ЕСЛИ у пациента высокая температура, ТО вероятность того, что у него ОРВИ, следует умножить на 5. Кроме правила продукций используются деревья решений, семантические сети и исчисление предикатов.

Машина вывода. Машина вывода – это высокоуровневый интерпретатор, который осуществляет цепочку рассуждений на основе фактов и правил базы знаний, и который приводит к конечному решению Машина вывода обычно имеет дело с ненадежными знаниями. Одна из проблем – работа с ненадежной информацией. В настоящее время найдены способы решения этой задачи: нечеткая логика, байесовская логика, коэффициенты уверенности. Эти способы дают на практике вполне приемлемые результаты.

Извлечение знаний. Извлечение знаний является трудоемким процессом. Знания сами по себе – дорогой ресурс, который сложно представить в простой для использования в компьютере форме. Обычный способ извлечения знаний состоит в том, что специалист по технологии ЭС опрашивает специалистов, знания которых добавляются в ЭС, добиваясь правильного представления их знаний в компьютере. Это

долгий и дорогой процесс. В настоящее время ведутся интенсивные работы по автоматизации процесса извлечения знаний. Появилось новое поколение систем -- самообучающиеся системы, которые уже нельзя назвать экспертными системами в точном понимании этого слова, т. к. они уже не используют знания экспертов. Процесс принятия решения в таких системах трудно понять человеку (не удается построить блок объяснения решения). Сейчас интенсивно развиваются системы, основанные на технологии нейронных сетей, которые используют этот принцип.

Система объяснения принятых решений. Человеку часто трудно объяснить, каким образом он принимает свое решение. Система объяснения принятых решений ЭС позволяет облегчить процесс общения человека с ЭС, объясняя, как система пришла к решению. В таком случае, при необходимости, человек может вмешаться в процесс принятия решения.

Области применения ЭС. С помощью ЭС эффективно решаются задачи, для которых нет строгой устоявшейся теории. Наиболее эффективное применение ЭС получили в диагностике, не только медицинской, но и технической и экономической. Приведем примеры медицинских экспертных систем.

Система MYCIN была спроектирована для ассистирования врачам в диагностике и назначении курса лечения при инфекционных церебральных заболеваниях. Медицинские знания о заболеваниях представлялись в виде продукционных правил "ЕСЛИ – ТО" в сочетании с коэффициентами неопределенности. Всего в систему было включено около 500 правил.

В системе CASNET (диагноз и прогнозирование течения глаукомы) рассматриваются ассоциативные связи между булевыми комбинациями результатов различных исследований и единственным патофизиологическим состоянием, которое оценивается с некоторым коэффициентом доверия.

В CASNET используются правила, связывающие с заданным правдоподобием эффект *sj* , следующий в результате появления причины *si*

. Кроме того, были предложены правила классификации, позволяющие по булевым комбинациям подтверждения или неподтверждения состояний указать диагностическое утверждение.

Система EXPERT позволяет задать продукционные правила с определенным уровнем доверия и иерархию этих правил. Эта система используется в ревматологии, эндокринологии и офтальмологии.

Система INTERNIST предназначена для диагностики внутренних заболеваний (более 500 заболеваний). В этой системе используются взаимосвязи двух видов: симптом – заболевание (частота встречаемости) и заболевание – симптом (интенсивность проявления).

MYCIN, EXPERT, CASNET являются логическими системами в том смысле, что взаимосвязи между всеми видами наблюдений за пациентами формируют основу вывода "ЕСЛИ – ТО", но сам вывод не обязательно имеет строго определенный характер, который связан с коэффициентами доверия или коэффициентами неопределенности.

Перечисленные выше системы считаются в настоящее время классическими в своей области и применяются в учебных целях.

Медицинская система ILIADF AND ILIAD ACCESS – ЭС, предназначенная для обучения и поддержки врача. Работает в двух режимах: как эксперт-консультант, предлагающий дифференциальный диагноз, и как модель пациента для обучения и тестирования (число моделируемых случаев заболевания потенциально не ограничено). База знаний системы содержит сведения более чем о 1200 внутренних расстройствах и приблизительно 6000 признаков (проявлений) болезней, что покрывает большинство известных заболеваний в следующих областях: кардиология, пульмонология, гастроэнтерология, эндокринология, неврология, инфекционные болезни, ревматология и гематология.

CLINDERM DERMAYOLOGIC DIFFERENTIAL DIAGNOSIS – аналитическая интерактивная система дифференциальной диагностики заболеваний кожи. Диагноз ставится на основе комбинации выраженных

морфологических аномалий и их расположения на теле. Программа предлагает список признаков для выявления их у пациента. Она может быть использована и для проверки предполагаемого диагноза: для любого заданного диагноза она выдаст дифференциальный диагноз в различных контекстах и описание возможных признаков заболевания.

CAPTURE – ЭС, служащая для регистрации случаев обнаружения рака и организации диспансерного учета больных в соответствии с рекомендациями Американского общества раковых заболеваний: автоматически следит за прохождением необходимых обследований, рассылает напоминания больным, хранит результаты анализов и устанавливает сроки последующих обследований.

Система MedTrac поддерживает ведение амбулаторной карты больного, используя базу знаний ЭС, охватывающую более 3000 диагнозов. Система осуществляет диагностический анализ и анализ риска по картам больных. Результаты лабораторных исследований заносятся автоматически из системы лабораторной поддержки ASTM-1238. SQL-поиск позволяет находить в базе данных группы больных с одинаковыми комбинациями симптомов и другую информацию. Для автоматического порождения текстов по картам система пользуется как медицинской терминологией, так и общеупотребительной лексикой. Кроме того, MedTrac автоматизирует выписку рецептов, памяток для пациентов, отслеживает коды диагнозов, необходимые для оплаты лечения и статистики, поддерживает интерфейс с головной базой данных.

К недостаткам большинства ЭС относятся их узкая специализация, отсутствие режима on-line работы с медицинским оборудованием, отсутствие интерфейса с лабораторными и офисными системами и между самими ЭС.

Экскурс по ЭС не является законченным без более подробного обсуждения ЭС Mycin. Mycin была одной из самых ранних ЭС, и ее проект сильно повлиял на дальнейшие проекты коммерческих экспертных систем и оболочек ЭС [77].

ЭС Mycin была разработана в Стенфордском университете в 1970-ых годах. Задачей системы является диагностика и лечение определенных классов инфекционных заболеваний крови. Диагностирование инфекционных заболеваний "обычным образом" включает в себя разведение культуры вируса. К сожалению, на это требовалось около 48 часов и если врачам приходилось ждать окончания этой процедуры, их больной мог умереть! Так что врачам нужно было быстро найти предположения относительно вероятных причин заболевания по имеющимся данным и использовать эти предположения для подбора лекарств.

Mycin был разработан частично с целью исследования постановки диагнозов экспертами на основании приблизительных (но важных) предположений, основанных на частичной информации. Однако, потенциально важной для практики была следующая проблема - имеется большое количество молодых докторов или докторов, не являющихся специалистами в данной области, которым иногда приходится ставить такие диагнозы, и если бы имелся какой-либо инструмент, который мог помочь им, тогда это могло бы позволить более эффективно проводить лечение. Фактически, Mycin никогда не использовался на практике. Это произошло не из-за недостатков в его эффективности - на тестировании он превзошел по быстродействию членов Стенфордской медицинской школы. Это произошло во многом из-за этических и юридических проблем, связанных с использованием компьютеров в медицине - если произойдет ошибка, кому предъявлять иск?

Mycin представлял знания в виде набора IF-THEN правил с коэффициентами уверенности. Правила были записаны на естественном английском языке. Вот пример одного из правил Mycin'a:

IF инфекция - первичная бактерия AND материал для посева культуры был отобран из стерильного источника AND предполагаемый путь внедрения микроорганизма - кишечный тракт THEN есть основания

предполагать (0.7), что этот микроорганизм имеет бактериальный характер.

Число 0.7 говорит о том, что степень уверенности весьма велика, но не носит абсолютный характер. Данные, подтверждающие гипотезу, накапливаются отдельно от данных, противоречащих ей, и "истинность" гипотезы в каждый момент времени есть алгебраическая сумма всех данных. Таким образом достигается комбинирование данных как в пользу, так и против одной и той же гипотезы.

Mycin был написан на языке программирования ЛИСП, и его правила формально представлены как выражения языка ЛИСП. Часть действия правила могла быть только заключением относительно решаемой проблемы, или она могла быть произвольным выражением. Это обеспечивало большую гибкость, но система проигрывала в модульности и ясности, присущих системам, основанным на правилах.

Метод решения, используемый в системе Mycin, представляет собой рассуждения в обратном направлении от диагностических гипотез к имеющимся данным, которое направляется правилами вывода. Однако для управления поиском решения (или для доказательства некоторой гипотезы) Mycin использовал различные эвристические знания. Они были необходимы, чтобы сделать рассуждения более эффективными и предотвратить задание пользователю ненужных вопросов.

Одна из стратегий состоит в том, что предварительно задаются пользователю более или менее общие вопросы, которые позволяют системе исключить маловероятные диагнозы. Затем система может сосредоточиться на частностях и использовать в полной мере режим вывода от цели к фактам для доказательства выдвигаемых гипотез.

Другие стратегии касаются способов вызова правил. Первая из них проста: используется первое возможное правило, Mycin первым делом проверяет все заключения правила, чтобы увидеть, является ли какой-либо из них заведомо ложным. Если это не так, то правило используется. Другая стратегия имеет больше отношение к коэффициентам уверенности.

Mycin сначала рассматривает правила, которые имеют более высокий коэффициент уверенности при выводе заключения, и правило перестает использоваться, как только суммарный коэффициент уверенности станет меньше 0.2.

Диалог с Mycin – сложный и длительный. Имеются три главных стадии диалога. В первой стадии собираются исходные сведения относительно случая, так что система может придумать очень большое число диагнозов. На второй стадии для проверки отдельных гипотез задаются более направленные вопросы. В конце этой стадии предлагается диагноз. На третьей стадии задаются вопросы для определения средств лечения на основании данных диагноза и фактов, относящихся к особенностям больного. Сеанс заканчивается рекомендациями относительно способов лечения. В любом месте пользователь может спросить, почему задан тот или иной вопрос или как было получено заключение.

Ситема Mycin – пионер в исследовании ЭС, в ней содержался ряд проблем, которые были исправлены в более поздних сложных архитектурах. Одна из проблем в том, что в базе знаний содержались вместе как знания о проблемной области, так и знания о способах решения задач и "условия просеивания правил", позволяющие не задавать пользователю глупых вопросов.

Более поздняя версия была названа NEOMYCIN, она пыталась использовать явную противоэпидемическую таксонометрию (представленную в виде фреймов) для представления фактов относительно различных видов эпидемий. Основная стратегия поиска решения задач состояла в том, чтобы спуститься на противоэпидемическом дереве от общих классов эпидемий к специфическим на основе собранной информация, чтобы различить два противоэпидемические подкласса (то есть, если болезнь 1 имеет подтипы болезней 2 и 3, и известно, что больной имеет болезнь 1, и подтип болезни 2 , а не 3 и имеет симптом 1, то следует задавать вопросы относительно симптома 1).

Имелись и другие разработки на основе проекта MYCIN. Например, EMYCIN была первой оболочкой ЭС, разработанной на основе Mycin. Новая ЭС, названная PUFF, была разработана для использования EMYCIN в области диагностики заболеваний сердца. Система, названная NEOMYCIN, разработана для обучения врачей на различных примерах заболеваний для проверки их заключений. Система могла объяснять ошибки в постановке диагноза. Необходимо пояснить, что не все экспертные системы подобны Mycin. Многие используют другие подходы к представлению знания и решению задач.

Наибольший интерес для практического здравоохранения представляют системы для диагностики и дифференциальной диагностики заболеваний. При этом для принятия решений могут использоваться самые разнообразные данные – анамнез, клинический осмотр (создаются ЭС диагностики, ограничивающиеся только этим набором), результаты лабораторных тестов и сложных функциональных методов. Список областей медицины, в которых начали применяться ЭС, обширен и продолжает расти.

4.2. Алгоритмическое и программное обеспечение экспертной системы диагностики оториноларингологических заболеваний

По данным различных эпидемиологических исследований, проведенных за последние 5 лет в более чем 30 странах, заболеваемость оториноларингологическими заболеваниями увеличилась в 2 раза, а удельный вес госпитализированных возрастает ежегодно на 1,5-2%. Поэтому в настоящее время лечение таких больных остается одной из самых насущных проблем здравоохранения.

В настоящее время лечение воспалительных заболеваний околоносовых пазух остается одной из самых злободневных задач оториноларингологии из-за крайней важности носового дыхания, т.к. оно обусловливает многие аспекты взаимоотношения организма человека с

окружающей средой. Важнейшей функцией носа является проведение воздуха в нижние отделы дыхательного тракта. Неадекватное носовое дыхание и раздражение слизистой оболочки полости носа приводят к целому ряду функциональных нарушений и развитию патологических процессов в органах и тканях. Топографические взаимоотношения носа и околоносовых пазух с полостью черепа и глазницы, единая кровеносная и лимфатическая сеть обусловливают сравнительно частый переход ряда патологических процессов, главным образом воспалительного характера, на полость черепа и особенно глазницы, вызывая тяжелые осложнения, порой приводящие к инвалидизации и угрожающие жизни больного [78, 79].

В повседневной клинической практике при постановке диагноза, особенно при обработке данных клинических и функциональных исследований, врачи до настоящего времени, в основном, используют собственный профессиональный опыт. Данные обстоятельства приводят к значительным затратам времени, субъективной оценке результатов и, как следствие, запоздалой постановке диагноза и сомнительности в его правдоподобии. Это относится и к оториноларингологии.

Проблемы оториноларингологических заболеваний и их осложнений имеют важное социальное значение, но они до настоящего времени остаются нерешёнными на должном уровне. Поэтому непосредственно наблюдаемые признаки, лабораторные и функциональные исследования, включая особенности связей между симптомами, характерные для атипичных вариантов заболеваний и редких нозологических форм, реализуемые в медицинских экспертных системах, могут значительно повысить качество диагностики, поскольку содержат в себе дополнительное врачебное знание.

В связи с этим создание ЭС оториноларингологических заболеваний [80] является важным для детей и взрослых. Система на основе Международного классификатора болезней МКБ-10 представляет методы и программное обеспечение, позволяющие наиболее эффективно,

экстренно и при минимуме дополнительных затрат комплексно оценить состояние больных с отоларингологическими заболеваниями.

ЭС реализована авторами монографии с использованием системы управления базами данных MS Access в среде языка программирования высокого уровня Delphi. Активизация каждой подсистемы осуществляется посредством выбора соответствующего пункта главного меню (рис.4.1).

Рис.4.1. Главное окно системы

Подсистема ввода данных принимает участие в работе системы как в режиме приобретения знаний, так и в режиме решения задач диагностики.

Основное назначение подсистемы ввода данных – заполнение базы данных, которая содержит данные анамнеза и диагноз пациента. При заполнении исходной базы данных, на основе которой вырабатывается база знаний, диагноз пациента устанавливает эксперт по диагностике.

В режиме решения задачи диагностики диагноз устанавливается ЭС на основе текущей базы знаний, но окончательное решение (вплоть до отмены установленного системой диагноза) принимает врач-оториноголог.

Ввод данных пациента осуществляется при выборе пункта главного меню посредством электронной анкеты (рис.4.2).

Рис. 4.2. Электронная анкета

Электронная анкета содержит пять закладок: "Общие сведения", "Жалобы пациента", "Заболевания ранее", " Провоцирующие, ухудшающие факторы", "Лабораторные исследования".

Каждая закладка содержит тематически объединенные симптомы, которые находятся в одном из двух состояний: в состоянии присутствия или в состоянии отсутствия и неопределенности (рис.4.3).

Рис. 4.3. Возможные значения признаков

После того, как на соответствующих вкладках заполнены сведения анамнеза, системой выдается предварительный диагноз (рис.4.4).

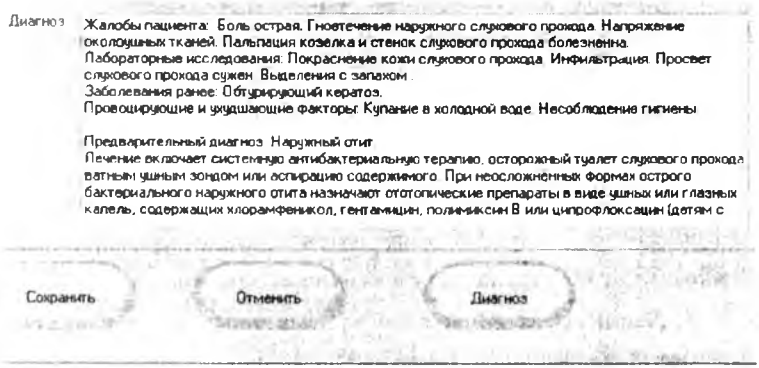


Рис.4.4. Предварительный диагноз, выданный системой

Сведения, вводимые посредством формы ввода данных пациента, хранятся в таблицах Obshie, Xarakteristika, Jalobi, Faktori, Lablssled, Lokalizacia, Prodoljitelnost, структура и связи между которыми представлены на рис.4.5.

Рис. 4.5. Структура базы данных

В заключение отметим преимущества созданной ЭС оториноларингологических заболеваний:

1. Стандартная электронная анкета и электронная база данных, облегчающие сбор, хранение и просмотр наиболее полной информации о пациентах.

2. Унификация данных и редактирование в случае неполных данных.

3. Легкость обновления и уточнения базы знаний.

4. Автоматизированная система приобретения знаний, предназначенная для извлечения интуитивных и трудноформализуемых знаний эксперта оториноларингологических заболеваний.

5. Диагностические правила как для типичных, так и для редких случаев оториноларингологических заболеваний.

6. Развертка возможных картин заболевания для данного диагноза.

7. Обследование картины заболевания и выделение существенных специфических параметров для различных подгрупп пациентов.

ГЛАВА 5. СПОСОБЫ АНАЛИЗА И ОБРАБОТКИ МЕДИКО-БИОЛОГИЧЕСКИХ ДАННЫХ

В последние годы начали широко применяться балльные системы оценки состояния больного, что объясняется повышенным желанием клиницистов прогнозировать результаты и лучше планировать интенсивную терапию [81]. В клинической практике наиболее популярны системы оценки:

1. Анестезия: индекс Goldman; шкала Mallam-patti; шкала Cormack и Lehane;

2. Интенсивная терапия: APACHE II и III; TISS (система оценки интенсивности лечебных вмешательств);

3. Травма: оценка травмы (TS); оценка тяжести травмы (ISS); педиатрическая оценка травмы (PTS);

4. Неврология: шкала комы Glasgow;

5. Педиатрия: оценка по Апгар.

В странах Западной Европы и США широкие эпидемиологические исследования показывают, что более 50% людей страдают повышенной кислотностью. Нарушения кислотно-шелочного состояния (КЩС) являются в большинстве случаев следствием серьезного патологического нарушения и редко имеют самостоятельное значение. Анализ КЩС позволяет отслеживать течение основного заболевания и контролировать эффект проводимой терапии. Исследование газового состава крови – незаменимый метод диагностики у пациентов с подозрением на респираторную патологию или метаболические нарушения.

5.1. Системы оценки состояния больного

Проблема объективной оценки тяжести состояния пациентов, находящихся на лечении в отделениях интенсивной терапии, актуальна в связи с необходимостью рандомизации групп пациентов для получения достоверных и сопоставимых данных. Объективизация состояния больных

позволяет прогнозировать вероятность летального исхода, сроки общей госпитализации и пребывания в отделении интенсивной терапии, оценивать эффективность проводимой терапии [82]. Необходимость объективизации тяжести состояния больных поддерживается многими учеными.

Современные системы объективной оценки тяжести состояния пациентов на основании методов, лежащих в их основе, можно классифицировать [83]:

1. Системы, оценивающие тяжесть состояния больного по количеству и сложности методов исследования и лечения, необходимых для его ведения (TISS);

2. Методы, основанные на статистическом моделировании (MPM);

3. Методы, основанные на сборе объективных клинических показателей и лабораторных данных (APACHE, SAPS).

Рассмотрим некоторые из них.

Индекс риска Goldman.

1. *Цель оценки*: прогнозирование периоперационной сердечно-сосудистой заболеваемости при некардиохирургическом вмешательстве путем определения факторов риска. Точность прогноза значительно варьируется в исследованиях различных клиник и у разных категорий больных в зависимости от используемых методов интенсивной терапии. Для более точной оценки степени риска необходимо использовать сравнительные данные в пределах отдельного объекта.

2. *Факторы риска*: ритм галопа, повышенное ЦВД, инфаркт миокарда в последние 6 мес, желудочковая экстрасистолия (более 5/мин), предсердные нарушения ритма, возраст более 70 лет, экстренные хирургические вмешательства, тяжелый аортальный склероз, тяжелое общее состояние больного, операции на органах грудной и брюшной полостей.

3. *Значение*. Каждый фактор риска оценивается по балльной системе, и результат суммируется. Если он превышает 26, то терапевтическое

вмешательство представляет повышенный риск в отношении осложнений на почве сердечной патологии. Доказано, что широкое применение индекса риска Goldman нецелесообразно ввиду межгоспитальных различий, изменений в лечебной тактике, специфики хирургического вмешательства и влияния сопутствующей патологии. Одним словом, анестезиолог или хирург может индивидуально оценить лишь абсолютный риск изолированного назначения.

Шкала APACHE.

1. *Цель.* Объект исследований — ряд клинических случаев, на основании оценки которых формируется прогностический результат лечения для группы реанимационных больных. Исследование не предназначено для индивидуальной оценки качества лечения или прогноза у отдельного пациента. До начала лечения определяется риск смертельного исхода в соответствии с имеющимся заболеванием и физиологическими резервами (в зависимости от возраста и наличия сопутствующих заболеваний), а также оценивается степень тяжести заболевания по выраженности расстройств физиологических функций. Количество переменных признаков сокращено с 33 (APACHE I) до 12 в APACHE II в связи с низкой прогностической значимостью. Система APACHE II представляет собой сумму балльных оценок настоящего физиологического состояния пациента и анамнестических данных, оценку сопутствующей патологии, а также степени срочности хирургического вмешательства. Стандартизированный показатель смертности (СПС) представляет отношение показателя настоящей смертности к прогнозируемой величине и может быть использован при анализе качества лечебной помощи.

2. *Оценка сопутствующей патологии.* По шкале оцениваются заболевания печени, почек и иммунологического статуса.

3. *Оценка настоящего физиологического состояния.* Для этого используются максимальные отклонения от физиологической нормы по 12 переменным, включая шкалу комы Glasgow, в течение первых 24 ч пребывания в ОИТ.

4. *Значение*. Шкала APACHE предназначена главным образом для оценки качества лечебной помощи. Она позволяет прогнозировать исход лечения определенных категорий больных, но не отдельно взятого пациента. Система не должна использоваться для составления прогноза в отношении летального исхода у одного больного. Решающее значение для точности прогноза имеет клиника заболевания.

APACHE III пытается улучшить систему оценки следующим образом: а) путем включения шести дополнительных переменных; б) посредством оценки сопутствующей патологии, функционального состояния и относительной смертности; в) путем увеличения количества диагностических категорий с 42 до 230 для улучшения прогнозирования.

Система оценки интенсивности лечебных вмешательств.

1. *Цель*: оценка тяжести состояния пациента по объему необходимой для него лечебной помощи. Каждое терапевтическое вмешательство, будь то вентиляция легких, физиотерапия или мониторинг артериального давления, оценивается баллами от 1 до 4, а затем результат суммируется.

2. *Значение*. Она позволяет определить степень тяжести заболевания и оценить возможность оказания необходимого объема лечебной помощи данному пациенту. Сравнение оценок по этой системе невозможно из-за специфичности лечебной помощи в различных ОИТ.

Оценка травмы.

Цель: определить состояние пациента при отборе и спрогнозировать отношение "гарантия качества — исход". Оценка основана на определении 5 параметров: а) число дыханий в 1 мин; б) напряженность дыхания; в) систолическое кровяное давление; г) наполнение капилляров; д) шкала комы Glasgow.

Значение. При оценке менее 12 баллов пациент должен госпитализироваться в травматологический центр. Показатель выживаемости при тупой травме или проникающей травме тесно связан с оценкой травмы в баллах.

Модифицированная оценка травмы.

Оцениваются только шкала Glasgow, систолическое артериальное давление и число дыханий в 1 мин. Каждый параметр кодируется и оценивается, показатели суммируются; общее значение варьирует в пределах от 0 до 7, 8408 баллов. Эта оценка имеет большое прогностическое значение, но не может использоваться при отборе больных.

Оценка тяжести травмы (ОТТ).

1. *Цель*: оценить сочетание повреждений по укороченной шкале травмы (УШТ).

2. *Методика.* Каждое из повреждений оценивается в зависимости от тяжести в баллах от 1 до 5. Повреждения кодируются и распределяются по одному на шесть зон: голова—шея, лицо, грудная клетка, живот, тазовые органы, кожные покровы или ожоги. ОТТ представляет собой сумму квадратов максимальных показателей, оцененных в 3 из 6 зон по УШТ. Максимальная оценка — 75 ($5\times5 + 5\times5 + 5\times5$).

3. *Значение.* Оценка тяжести травмы значительно коррелирует с показателями заболеваемости, смертности и длительности пребывания больного в стационаре. Она применяется при обследовании пациентов старше 12 лет.

Оценка травмы у детей.

1. *Цель.* Оценочный отбор детей рекомендуется осуществлять по системе опережающего поддержания жизни при травме. Оценки определяются от -1 до +2 в зависимости от возраста ребенка, проходимости дыхательных путей, систолического артериального давления, уровня сознания, типа переломов и обширности повреждений покровных тканей.

2. *Значение.* Оценка травмы ниже 8 баллов указывает на повышенный риск заболевания и смерти, а также служит показанием к госпитализации в травматологический центр.

Оценка по Апгар.

1. *Цель*: определить состояние новорожденного по простой оценочной шкале. Оценки от 0 до 2 отражают характер показателей (частота сердечных сокращений, напряжение дыхания, мышечный тонус, двигательные рефлексы и цвет кожных покровов).

2. *Значение*. Она широко используется для оценки состояния ребенка при рождении; однако в качестве прогностического критерия она не показательна. В этом плане более важное значение имеет оценка неврологического статуса.

Шкала комы Glasgow.

1. *Цель*: определить глубину комы путем оценки вербальных, моторных и глазных реакций на стимуляцию. Открывание глаз может быть спонтанным (4), по просьбе (3), в ответ на боль (2) или отсутствует (1). Состояние двигательных реакций определяется: согласно командам (6); направленностью к болевому раздражителю (5); при болевом раздражении (4); при патологическом сгибании (3); разгибании (2); отсутствии реакций (1). Вербальные реакции оцениваются следующим образом: ориентированные (5); спутанные (4); не соответствующие словам (3); на все звуки (2); отсутствие (1).

2. *Значение*. Общая оценка распределяется от 3 до 15. Данные, получаемые по этой шкале, позволяют определить индекс глубины комы, имеют некоторое прогностическое значение и сравнимы с результатами, получаемыми в других отделениях и клиниках. При оценке ниже 8 показана вентиляция легких. Шкала требует уточнения, так как аналогичные оценки могут быть получены при различных коматозных состояниях

Все системы независимо от количества входящих в них параметров по таким критериям, как чувствительность, специфичность, общая корректность, находятся приблизительно на одном уровне и могут быть использованы в клинических условиях с одинаковой эффективностью.

5.2. Алгоритмическое и программное обеспечение оценки тяжести состояния пациентов

Различие показателей летальности обусловлено целым рядом объективных и субъективных причин, не зависящих от уровня оказания медицинской помощи, а определенных неоднозначным подходом к их измерению. Объективизация оценки тяжести состояния пациентов невозможна без использования многоступенчатых шкал оценки тяжести. Стратификация больных в соответствии со шкалами позволит оптимизировать процесс статистического анализа, повысить его достоверность и объективность и тем самым улучшить эффективность лечебных мероприятий в сравнимых группах пациентов.

Как указывалось в 5.1, наиболее информативными для оценки тяжести состояния пациентов с АС можно считать шкалы SIRS, SAPS, SOFA и MODS. Быстро оценивать статус больного либо пострадавшего возможно по шкалам APACHE, SAPS, CRIB, Глазго, АПГАР.

Прогнозирование результатов лечения даёт возможность объективного выбора лечебной тактики, оценки эффективности и экономического обоснования целесообразности того или иного метода терапии. Одним из инструментов современного прогнозирования являются разработанные в результате сложного математического анализа интегральные шкалы оценки тяжести состояния (APACHE II, SAPS II, шкала комы Глазго - GCS, SOFA и MODS), которые в течение последних лет используются в отделениях реанимации различного профиля.

Развитие современной медицины невозможно без внедрения в клиническую практику процесса прогнозирования. Опережающее отражение результатов интенсивного лечения особенно значимо для пациентов с тяжёлой травмой. Данное обстоятельство связано с увеличением числа пострадавших, высокой летальностью и значительной степенью их инвалидизации. Так, по данным ВОЗ ежегодно от травм погибает до 2 млн. человек. Во всём мире у мужчин в возрасте до 45 лет и

у женщин до 35 лет травматические повреждения - главная причина смерти [83, 84].

Оценивая ущерб от тяжелой травмы, необходимо отметить, что по количеству непрожитых лет ущерб от травм значительно превышает таковой от сердечно-сосудистых, онкологических и инфекционных заболеваний вместе взятых.

Расчетные программы для врачей отделений интенсивной терапии, анестезиологии и реанимации входят в медицинское программное обеспечение. Они направлены на оптимизацию труда медицинских работников и разрабатываются с целью обучения, для практической работы, диагностики, прогнозирования и проведения научных исследований.

Разработанная авторами монографии программа-калькулятор позволяет быстро оценивать статус по шкалам APACHE, SAPS, CRIB, Глазго, АПГАР. Главное окно программы представляет вверху шкалы, в которых производятся расчеты, ниже – окна с лабораторно-клиническими данными, которые необходимо ввести для расчета, например, в шкале APACHE III (рис.5.1).

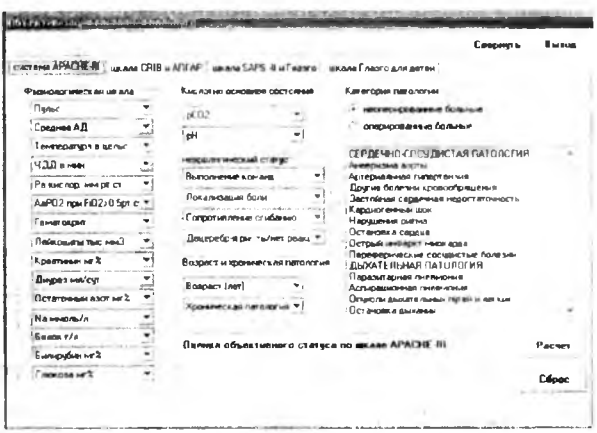


Рис.5.1.Интерфейс шкалы APACHE III

Переход для расчета в нужной шкале проводится в верхних окнах программы. Данные лабораторно-клинических исследований вводятся в соответствующие окна для каждой шкалы. Введя данные и нажав кнопку «Расчет», получим оценку состояния больного в баллах для нужной шкалы и прогноз (рис.5.2).

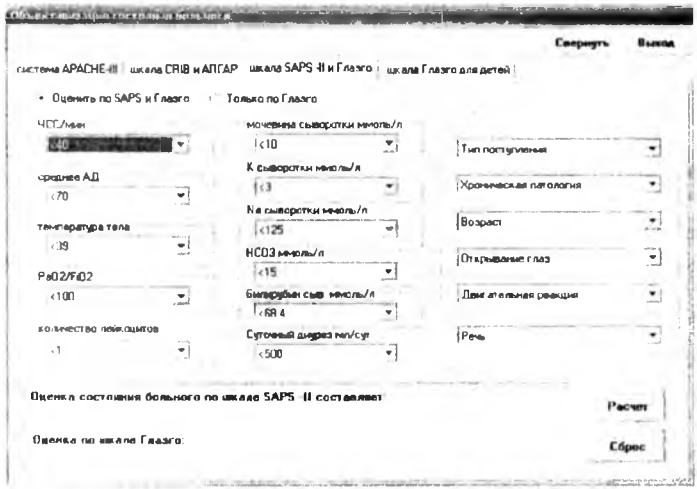


Рис.5.2. Интерфейс шкалы SAPS-II и Глазго

Особенно важно то, что объективизация состояния больного калькулятивно вычисляется в шкале APACHE III, которая, по нашему мнению, является наиболее полной, так как в категории неоперированных больных подробно перечислены патологии различных систем, а у оперированных пациентов указаны операция и диагноз.

Это позволяет быстро и точно провести объективизацию состояния больного и прогнозирование, особенно в узкоспециализированных отделениях: онко-, кардио-, нейрохирургических. Алгоритм расчетов взят из наиболее современной, актуальной и признанной литературы [84]. Интерфейс программы удобен и понятен, работа стабильна.

Вышеописанное программное обеспечение может найти широкое применение в клиниках как у практических врачей, так и у занимающихся научными исследованиями.

5.3. Автоматизация обработки данных при оценке кислотно-щелочного состояния организма

Анализ КЩС – важная часть изучения спонтанного или искусственного дыхания и процессов метаболизма в организме. Показатели КЩС в клинической и экспериментальной практике являются одними из важнейших, характеризующих степень постоянства внутренней среды организма, а также определяющих необходимый уровень ее коррекции [85]. В регуляции КЩС участвует несколько механизмов, к ним относятся буферные свойства крови, газообмен в легких и выделительная функция почек.

Для правильной оценки требуется определение таких показателей, как: pH крови, т. е. концентрация водородных ионов в крови, характеризующая ее кислотность или щелочность; Pco_2 – парциальное давление (напряжение) углекислого газа в крови (в мм рт. ст.), HCO_3 – содержание истинных бикарбонатов (ммоль/л); $[HCO_3]S$ – стандартный бикарбонат (ммоль/л); ΣCO_2 – общее содержание углекислоты (ммоль/л); BE – избыток (смещение) оснований (ммоль/л).

Методика и аппаратура для определения исходных показателей КЩС являются в достаточной степени отработанными и обеспечивающими необходимую точность результатов. Процесс измерения параметров КЩС достаточно сложен, сопряжен с рядом операций, автоматизация которых связана со значительными техническими трудностями [86].

Поэтому первоочередной задачей, от решения которой во многом зависит эффективность использования параметров КЩС во врачебной практике, и при проведении научных исследований, является не столько автоматизация собственно процесса измерения исходных показателей,

сколько автоматизированная их обработка и оперативное получение в конечном итоге полного перечня значений всех параметров КЩС.

Авторами монографии разработано программное обеспечение по оценке указанных выше параметров (рис.5.3).

Рис.5.3. Стартовая панель программы

Оно предусматривает определение исходных непосредственно измеряемых показателей с помощью обычных методик на стандартном оборудовании с последующим вводом результатов в компьютер (рис.5.4).

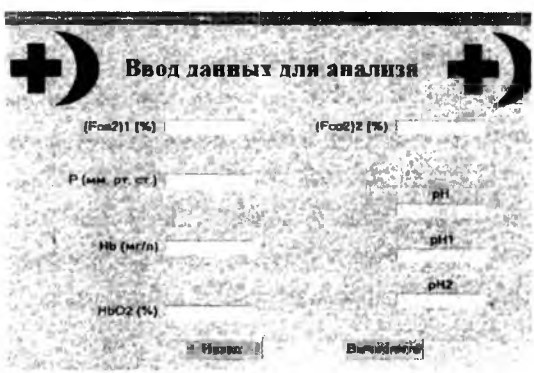


Рис.5.4. Модуль ввода экспериментально измеренных параметров КЩС

Затем программа осуществляет программную обработку вводимой информации и расчёт параметров КЩС (рис.5.5).

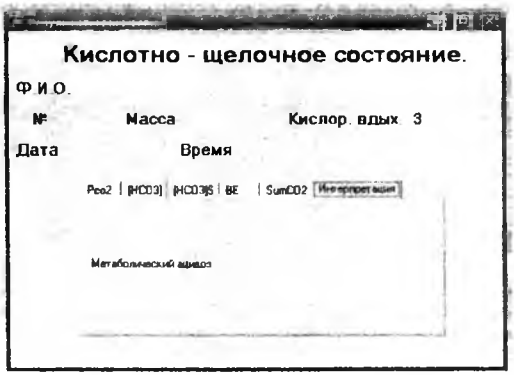


Рис.5.5. Интерпретация результатов для диагностики различного рода сдвигов КЩС

Для формирования диагностических заключений алгоритм сопоставляет найденные показатели КЩС с определенными граничными практическими значениями, которые в виде таблицы предварительно помещают в оперативную память. Диагностический вывод осуществляется по трем показателям КЩС – pH, Pco_2 и BE.

Оценка КЩС крови имеет большое значение в клинике. Она позволяет выявить у больного ацидоз либо алкалоз и судить о том, является он респираторным или нереспираторным. Острые, тяжелые заболевания и отравления, которые вызывают нарушение КЩС организма, как правило, компенсируются в условиях лечебных учреждений, так как опасны для жизни и имеют ярко выраженную симптоматику. Заключение о состоянии КЩС позволяет выбрать правильное лечение.

5.4. Автоматизированная обработка данных клинических функциональных исследований дыхания

Различные патологические нарушения работы сердца обусловливают, как правило, и легочные заболевания различной степени тяжести. В связи с этим необходимо решение задач, касающихся обработки количественных данных, характеризующих состояние и функцию лёгких и облегчающих диагностику, наблюдение течения заболевания и его лечение. Существующие эффективные методики сбора и обработки данных о функционировании системы дыхания значительно облегчают диагностику лёгочных заболеваний [87]. На основе обобщения и анализа опыта исследований физиологии дыхания авторами монографии совместно с врачами Республиканского специализированного центра хирургии им. академика В. Вахидова с целью повышения точности постановки диагноза заболеваний лёгких разработан программный комплекс (ПК) автоматизированной обработки данных исследования внешнего дыхания [88].

В созданном ПК обработка биомедицинских данных выполняется в графическом интерфейсе на основе инструментальных средств объектно-ориентированного языка PHP [89], разработанного для создания WEB-ориентированных приложений с хорошей программной переносимостью, и включает следующие основные этапы:

- ввод исходной информации (рис. 5.6);

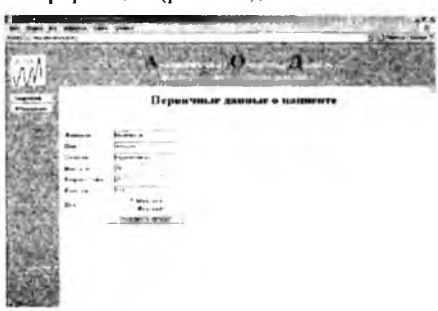


Рис.5.6. Первичные данные о пациенте

- ввод значений первичных характеристик внешнего дыхания, измеренных по стандартным методикам (максимальная вентиляция легких, максимальное количество воздуха, которое может быть провентилировано в 1 мин; минутное поглощение кислорода; объем форсированного выдоха за 1 с; форсированная жизненная емкость легких) (рис.5.7);

Рис.5.7. Первичный расчёт данных

- расчёт основных показателей дыхательной системы (рис.5.8);

Рис.5.8. Значения показателей дыхательного объёма

- расчёт должных значений показателей и соотношений между действительными и должными показателями (рис.5.9);

Рис.5.9. Значения основных показателей дыхательной системы

- постановка диагноза на основе интерпретации полученных результатов;

- формирование и выдача предварительного диагноза.

Значения первичных характеристик внешнего дыхания определяются на спирографе серийного производства (Spirosift 5000, фирма Fukuda Denshi, Япония) по стандартной методике и вводятся в ПК по мере получения. Порядок ввода данных происходит в строгом соблюдении последовательности. Обработка исходных данных и получение основных показателей дыхательной системы осуществляется по завершении всех процедур обследования в целом.

Разработанный программный диагностический комплекс предоставляет возможность установления предположительного диагноза или нормы. Кроме того, при проведении предварительного диагностирования в ПК предусмотрено решение задачи выявления дыхательной недостаточности, определение её степени и формы, то есть такого состояния организма, при котором нарушается нормальная функция дыхательного аппарата, обеспечивающая процессы газообмена. Сопоставление граничных значений ряда показателей внешнего дыхания с фактическими, полученными по данным спирографии, позволяет оценить степень дыхательной недостаточности. Аналогичным образом определяется и её форма.

В результате проведения исследования с помощью ПК выдаётся заключение, в которое включаются рассчитанные параметры (первичные и вторичные) и интерпретация полученных результатов (предварительный диагноз), например, рестриктивная форма вентиляционной недостаточности. По необходимости ПК можно легко дополнить и изменить форму результатов введением в него новых показателей, имеющих значения для диагностики. Созданный ПК способствует эффективности и облегчению процесса диагностики. Окончательное решение в выборе диагноза - за лечащим врачом.

ГЛАВА 6. АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ВЕДЕНИЯ ИСТОРИИ БОЛЕЗНИ

Основной технологический процесс в здравоохранении – лечебно-диагностический. Медицина является информационно-зависимым видом деятельности. В ходе лечебно-диагностического процесса материальные потоки незначительны, а информационные – велики и сложно организованы. Участники лечебно-диагностического процесса передают друг другу большое количество сведений об объекте этого процесса – пациенте.

Именно сообщения о состоянии пациента и ходе его лечения, которыми в различных формах обмениваются между собой медики, организуют лечебно-диагностический процесс и обеспечивают его целостность. И в то же время персональные сведения о пациенте, фиксируемые в ходе выполнения лечебно-диагностических мероприятий, являются входными данными для систем, обеспечивающих административно-хозяйственную деятельность медицинского учреждения [90].

Можно сказать, что ведение истории болезни, или, выражаясь более точно, документирование лечебно-диагностического процесса, является базисным информационным процессом в любом ЛПУ.

Документирование лечебно-диагностического процесса не сводится к заполнению истории болезни, так как сведения о пациенте заносятся также во множество журналов и на множество бланков. В результате информатизации медицинского технологического процесса необходимо получать актуальные и достоверные данные, необходимые для принятия решений, поэтому создание электронной истории болезни (ЭИБ) должно быть одним из первых шагов на этом пути [91].

6.1. Основные факторы ведения истории болезни

История болезни – это сжатый отфильтрованный отчет об эпизодах лечения пациента в системе здравоохранения.

Цели ведения истории болезни можно подразделить на три группы: ведение истории болезни способствует лечению пациента, обеспечивает финансовую и юридическую отчетность и помогает проведению клинических исследований. Поскольку история болезни является делом рук человека, то цели ее ведения далеко не являются непреложными. Можно ожидать, что функции истории болезни будут изменяться по мере того, как новые технологии обеспечат альтернативные методы регистрации и анализа данных, а финансовые и юридические органы установят новые требования к ведению документации и отчетности.

Основной целью ведения истории болезни является содействие лечению пациента. История болезни обобщает то, что было с пациентом в прошлом, и документирует наблюдения, диагностические выводы и планы медицинского персонала. В определенном смысле она является внешней памятью, к которой специалисты здравоохранения могут обратиться, когда они вспомнят о пациенте спустя некоторое время.

История болезни является также средством взаимодействия между специалистами и обратившимися к ним врачами, между врачами и медицинскими сестрами и т.д. В больнице она является основным проводником действий. Врачи инициируют диагностические и терапевтические действия, записывая соответствующие распоряжения на бланках рецептов и заказов (направлений). Сотрудники, получающие рецепты и заказы, в свою очередь записывают свои действия и наблюдения; например, лаборанты записывают результаты лабораторных тестов, фармацевты регистрируют отпуск лекарств, а медицинские сестры записывают детали своего взаимодействия с пациентами [92].

Больничная история болезни является основным механизмом, обеспечивающим преемственность лечения в течение госпитализации

пациента. Амбулаторная история болезни помогает обеспечить преемственность лечения от одного визита пациента к другому. Поскольку ожидаемая продолжительность жизни растет и популяция стареет, центр тяжести амбулаторного медицинского обслуживания смещается в сторону профилактики и лечения хронических заболеваний, а не лечения острых заболеваний. Амбулаторная история болезни позволяет медицинским работникам просматривать данные, собранные за достаточно большие промежутки времени, и тем самым изучать течение проблем и заболеваний пациента.

Типичным недостатком бумажной истории болезни является ее недоступность. В больших больницах традиционные истории болезни могут оказаться недоступными в течение нескольких дней из-за того, что они используются в административном офисе либо сложены в ожидании, пока лечащий врач не сделает выписной эпикриз. Если информация из истории болезни хранится в компьютере, то при наличии доступа к терминалу компьютера врач может получить эту информацию за несколько секунд, вместо того, чтобы ждать минуты или часы, необходимые для поиска и доставки бумажной истории болезни.

Хранение записей в памяти компьютера позволяет обеспечить к ним удаленный доступ, например, врач может просматривать их, находясь у себя дома. Оно позволяет также одновременный доступ; например, в одной комнате медицинская сестра может просматривать динамику изменения артериального давления у данного пациента, а в другом помещении врач может анализировать результаты выполненных для этого же пациента лабораторных тестов – ситуация, совершенно невозможная при наличии только бумажной истории болезни.

Автоматизированные системы ведения истории болезни обеспечивают предоставление более разборчивых и лучше организованных отчетов. Улучшение разборчивости связано с тем, что отчеты печатаются, а не составляются от руки, а лучшая организация есть следствие того, что компьютеры создают структуру хранящимся в них данных. Компьютеры

могут обеспечить повышение полноты и качества введенных данных за счет автоматически выполняемых проверок. Более того, диалоговые системы могут запрашивать у пользователя дополнительную информацию – свойство, которое неспособна обеспечить ни одна бумажная форма статистического учета. Наконец, компьютеры могут способствовать процессу ввода данных и более сложными методами, например, путем управления потоком входных форм или с помощью проверок, что формирование требуемых отчетов завершено.

Медицинские записи, хранящиеся в памяти компьютера, могут предоставляться на разных носителях информации, начиная от экранов видеотерминалов до бумаги. Конечно, хранение медицинских записей в памяти компьютера вовсе не означает отказ от бумажных документов. Кроме того, при использовании компьютеров одни и те же данные могут быть представлены во многих формах; запись о визите пациента, ответ врачу, направившему пациента на консультацию, а также врачебное заключение могут содержать одну и ту же информацию. Форма и содержание отчета, выданного компьютером, могут быть приведены в соответствие назначению отчета – тем самым снижается избыточность затрат ручного труда на переписывание одних и тех же данных. Кроме того, информация о многих пациентах может быть агрегирована – полезное свойство как для ведения научной работы, так и для управления процессом лечения.

Хранение записей в памяти компьютера имеет и то большое преимущество, что компьютер может автоматически принимать решения о данных, которые он собирает и выдает. Как уже отмечалось ранее, система может запрашивать у пользователя важную отсутствующую информацию. Еще важнее то, что компьютер может анализировать данные и помогать медицинскому персоналу ставить диагнозы и принимать терапевтические решения.

Степень полноты реализации этих преимуществ в конкретной системе электронного ведения истории болезни зависит от следующих четырех факторов:

1. Спектр информации, охватываемый системой. Содержит ли система результаты, полученные в амбулаторных учреждениях или и в других учреждениях тоже? Содержит ли она только информацию о лекарственной терапии и лабораторных тестах, или еще и результаты выполненных врачами осмотров?

2. Продолжительность использования системы. Во многих ситуациях записи, аккумулирующие данные о пациенте за последние пять лет, будут более ценными, нежели записи о визитах пациента за один конкретный месяц.

3. Форма представления данных в системе. Медицинские данные могут храниться в повествовательной форме, и быть всего лишь более разборчивыми и доступными, нежели их бумажные эквиваленты. Однако некодированная информация не стандартизуется, и недостаточно последовательное применение медицинской терминологии снижает возможности поиска необходимых данных. Лишь в том случае, когда используется контролируемый заранее определенный словарь терминов, можно агрегировать и обобщать данные, предоставленные разными врачами или тем же самым врачом в разное время. Таким образом, запись неструктурированных данных не может достаточно активно способствовать принятию решений или проведению научных исследований.

4. Географическое распределение терминалов, обеспечивающих доступ к системе. Если большое число пользователей будут иметь доступ к системе только из ограниченного числа мест, то она будет менее ценной, чем аналогичная система, доступная с нескольких сотен терминалов, расставленных по всей больнице или даже за ее пределами, на дому у врачей или в кабинетах частнопрактикующих специалистов.

6.2. Фундаментальные вопросы разработки и внедрения автоматизированных систем ведения истории болезни

Цели у всех систем ведения истории болезни одинаковы независимо от того, какая технология применяется – ручная или автоматизированная. Однако механизмы достижения этих целей отличаются. С точки зрения пользователей, фундаментальное различие двух указанных подходов состоит в способах занесения сведений в историю болезни и выборки из нее необходимой информации [92].

Передача данных из источника их возникновения в компьютер требует выполнения двух отдельных процедур:

1. Получение информации;
2. Ввода данных.

Получение информации. Если все регистрируемые в истории болезни сведения порождаются в рамках одной медицинской организации, отвечающей за ведение истории болезни, то процедура получения информации является достаточно тривиальной. Врач может без особого труда получить информацию, собранную внутри учреждения, а также результаты диагностических исследований и лабораторных тестов, заказанных другим учреждениям.

С другой стороны, получение аналогичной информации, собранной при госпитализации пациента в другом учреждении, при оказании ему скорой и неотложной помощи, при посещении внешнего консультанта, может оказаться затруднительным или невозможным. Эта информация может оказаться пропущенной (например, пациент забыл при визите к врачу упомянуть о недавно прошедшей госпитализации), неразборчивой (скажем, на третьем экземпляре карты учета скорой и неотложной помощи невозможно что-либо прочитать), недостаточно детальной (к примеру, консультант сообщил, что все результаты исследования пациента в пределах нормы, но не указал сами результаты). Для разрешения этих

проблем приходится проводить переговоры с теми местами, откуда такая информация часто поступает, и выполнять дополнительную работу.

Может оказаться необходимым ограничить сферу ведения автоматизированной истории болезни только теми сведениями, которые возвращаются в данное учреждение, но это будет способствовать понижению возможности программного обеспечения компьютера по предоставлению полезной информации о лечении пациента. Аналогично автоматизированная система ведения амбулаторной истории болезни будет иметь ограниченные возможности выдавать предупреждения и напоминания, если данные, собранные в одном подразделении, не будут доступны другому подразделению. Современная тенденция создания больших, более интегрированных и более самодостаточных медицинских автоматизированных систем позволит рассчитывать на то, что со временем проблема получения информации станет менее острой.

Ввод данных. Информация о пациенте, собираемая врачами, требует специальных комментариев, поскольку ее обработка представляет собой наибольшие трудности для тех, кто разрабатывает и эксплуатирует автоматизированные системы ведения истории болезни. Врачи регистрируют четыре типа информации:

1. Анамнез, то есть сведения со слов пациента или его близких, например история заболевания или текущие симптомы;

2. Данные обследования пациента, проведенного врачом;

3. Дифференциальный диагноз, поставленный врачом;

4. План лечения пациента.

Некоторая информация (например, история заболевания, обычно заполняемая врачом) может быть получена другими способами, например, с помощью вопросника, из беседы пациента с медсестрой или путем диалога пациента с компьютером. Однако возможно, что данные об истории заболевания, собранные с помощью вопросников или через медсестер, гораздо менее продуктивны для установления диагноза, нежели аналогичные данные, собранные врачами. Следовательно, нельзя быть

уверенными, что данные эквивалентного вида, полученные разными способами, будут содержать эквивалентную информацию.

Записи врачей могут вводиться с помощью одного из трех способов: операторского ввода рукописных или надиктованных записей, ввода данных из заполненных врачами формализованных бланков либо непосредственного ввода данных самими врачами. Операторский ввод записей особенно уместен, если учреждение уже вложило средства в диктофонную технологию, поскольку в этом случае стоимость набора текста уже входит в затраты учреждения. Если при диктовке врачи следуют определенным стандартам изложения, то оператор может вводить данные в умеренно структурированном виде. Например, если врач диктует свою информацию, используя стандартный порядок (анамнез болезни, анамнез жизни, данные обследования и план лечения), то оператор может вводить каждую часть диктовки в отдельное поле экранной формы, изображаемой ему на дисплее.

Второй метод предполагает, что врачи записывают информацию о визите пациента на формализованных бланках, с которых затем данные вводятся (и, возможно, кодируются) вспомогательным персоналом. В настоящее время это один из самых успешных подходов.

Третьей альтернативой является ввод данных непосредственно врачами, используя видеотерминалы (дисплеи). В ряде больниц врачи сами вводят лекарственные назначения и заказы на лабораторные тесты. Непосредственный ввод заказов наиболее легко принимается хирургами, поскольку они могут создать небольшое число стандартных шаблонов заказов, пригодных для большинства своих пациентов, и вводить эти заказы с помощью нажатия нескольких клавиш, что обеспечивает значительное ускорение по сравнению с ручным оформлением. Терапевты и семейные врачи, которые лечат пациентов с гораздо большим разнообразием клинических проблем, менее склонны к непосредственному вводу, поскольку для ввода данных требуется больше времени, а

процедура ввода требует умения набирать на клавиатуре, чему многие врачи не хотят учиться.

Непосредственный ввод анамнеза, результатов осмотра и дневников еще менее приемлем для врачей, нежели ввод заказов, поскольку связан с продолжительным диалогом, необходимым для ввода этих данных в компьютер. Сопротивление непосредственному вводу данных в компьютер ослабевало с появлением рабочих станций на микрокомпьютерах, дисплеев с высокой разрешающей способностью, манипуляторов типа мыши и трекбола, а также речевого ввода. Последний способ ввода является наиболее многообещающим, поскольку врачи уже знакомы с технологией диктовки информации о пациентах и нередко предпочитают ее всем другим.

Уже существуют коммерчески доступные системы, позволяющие записать речь рентгенолога, диктующего свое заключение, и воспроизвести ее по телефону любому медицинскому работнику, которому это заключение может понадобиться быстро. Если компьютер обеспечивает ввод речи, то врачи могут кодировать наиболее часто встречающиеся предложения с помощью выбора из меню, а затем вводить остальную информацию, просто диктуя ее на компьютер. Распознавание речи компьютером является еще более многообещающим, поскольку оно может позволить компьютеру "понимать" устные команды и преобразовывать их в соответствующие коды или текст. Однако первые эксперименты по распознаванию речи, проведенные в клинических условиях, не привели к успеху из-за слишком высокого процента ошибок распознавания. Более поздние системы оказались более успешными. Система Курцвейля достигла надежности распознавания 95%, но при этом она еще не может обеспечить распознавание слитной речи. Кроме того, ее пользователи при диктовке своих текстов должны ограничиваться заранее заданным словарем.

Процедура ввода данных является трудоемкой и занимает у персонала довольно большое время. Люди должны интерпретировать данные или

перевести их в другую форму, а затем ввести в компьютер. Данные могут вводиться в виде свободного текста, в закодированном виде или в форме сочетания свободного текста и кодов процедур.

Основное преимущество кодирования данных состоит в том, что тем самым данные классифицируются и стандартизуются, а это облегчает ведение научной работы, формирование счетов на оплату лечения, а также последующую выборку историй болезни. Кодирование позволяет компьютеру "понимать" данные и выполнять более разумную обработку этих данных. Кроме того, для хранения закодированных данных обычно требуется гораздо меньше места, чем для некодированных; если же возможных кодов немного, то информацию можно вводить более удобным образом, выбирая соответствующие кодам строки меню.

Основным недостатком кодирования является стоимость преобразования исходных текстов в правильные коды. Для классификации исходного текста требуется определенное время, особенно если этот текст не вполне обычен и непосредственно не укладывается в один из существующих классов. На обучение персонала процессу кодирования также требуется время; ведение словаря кодов, описывающего соответствие между кодами и обозначенными ими терминами, тоже является трудоемким делом. При кодировании могут возникать ошибки, которые по сравнению с опечатками в свободном тексте гораздо труднее обнаружить, поскольку в закодированной информации отсутствует избыточность, присущая свободному тексту. Например, можно ввести код 392 вместо 329 и для компьютера это не будет ошибкой; он может только предложить оператору ассоциированный с кодом текст, для последующей визуальной проверки.

Между кодированием и вводом свободного текста существует определенное противопоставление. Чем больше используется кодирование, тем больше времени тратится на интерпретацию данных; чем больше используется свободный текст, тем больше времени тратится на ввод данных. Кодирование предпочтительнее в ситуациях, когда число

возможных кодов невелико или персонал, занимающийся вводом данных, имеет приемлемое медицинское образование и выполняет эту работу в течение времени, позволяющего усвоить и эффективно применять более сложные схемы классификации. Напротив, свободный текст, предпочтительнее, если число возможных кодов велико, а персонал, обеспечивающий ввод данных, не может быть достаточно быстро натренирован для выполнения сложной интерпретации данных. Система ведения истории болезни может сочетать оба метода, применяя кодирование для распространенных диагнозов и результатов исследований, а ввод свободным текстом - для остальной информации.

Кодирование рукописных заметок, сделанных врачами, представляет известную трудность в связи с неразборчивостью почерка.

Наличие в электронном виде закодированных данных о пациенте, передаваемых из лабораторных и аптечных автоматизированных систем, значительно упрощает ввод клинической информации в компьютер. Оно исключает необходимость набирать эти данные на клавиатуре и может уменьшить работу по кодированию, хотя обычно не исключает ее полностью. Кодирование может остаться проблемой, поскольку схема кодирования, использованная в системе, передавшей данные, например в лабораторной системе, может не вполне совпадать с той схемой, что применяется в автоматизированной системе ведения истории болезни. К примеру, одна из лабораторий может выбрать для своих результатов шкалу от 1 до 4, а другая ограничиваться градациями нормальный/аномальный. В различных системах даты, время и идентификаторы пациента нередко имеют несовместимые форматы. Поэтому персоналу, который пользуется автоматизированной системой ведения истории болезни, нередко приходится транслировать внешние кодированные данные во внутренние коды своей системы.

Основными препятствиями для широкого применения автоматизированных систем ведения истории болезни являются их высокая стоимость, задержки и возможные ошибки, присущие ручному

вводу данных. Эти препятствия могли бы быть преодолены, если бы данные получались на месте их возникновения сразу в машиночитаемом виде. Однако для того, чтобы такая технология стала возможной, необходимо обеспечить унификацию форматов данных и определенную степень стандартизации схем кодирования информации.

Предотвращение ошибок. Вследствие возможности возникновения ошибок при вводе клинической информации в компьютер автоматизированные системы ведения истории болезни должны выполнять тщательные проверки вводимых данных. При вводе клинических данных можно использовать целый ряд различных методов проверки. Проверки выхода за границы могут выявлять или предотвращать ввод данных, выходящих за пределы допустимых значений (например, концентрации калия в сыворотке крови 50,0 — нормальный диапазон значений концентрации для здорового человека составляет от 3,5 до 5,0 ммоль/л).

Проверки по шаблону могут анализировать соответствие вводимых данных определенному шаблону (например, три цифры, затем дефис и еще четыре цифры для городского телефонного номера). Численные проверки могут определять соответствие введенных данных определенной математической формуле (например, сумма значений распределения лейкоцитов, выражаемых в процентах, должна равняться 100). Проверки совместимости могут выявлять ошибки, сравнивая несколько введенных значений (например, обнаружить код рака простаты, введенный как диагноз для женщины). Проверки отклонений предупреждают о больших и необычных отклонениях новых данных от предыдущих (например, значение веса одного и того же пациента изменилось на 40 кг за 2 недели). Проверки орфографии сравнивают правильность написания отдельных слов.

Заключения и эпикризы. Автоматизированное ведение истории болезни позволяет представлять важные компоненты истории болезни в виде подборки компактных и более обозримых документов. Обычно для этого выбираются специфические классы данных о пациенте, например

активные аллергии, активные проблемы, активное лечение и результаты последних осмотров. Хорошим примером может служить заключительный эпикриз, сформированный системой COSTAR.

В будущем можно ожидать появления более сложных стратегий составления заключений и эпикризов, связанных, например, с выявлением значительных отклонений в параметрах состояния пациента или с агрегированием в одном диагностическом заключении аномальных значений параметров близкой природы (например, повышенного содержания трансаминазы (SGOT), повышенного содержания щелочной фосфатазы и билирубина, каждый из которых является индикатором нарушений функций печени).

Можно будет встретить заключения, в которых различаются аномальные значения параметров, на которые направлено лечение, от тех параметров, которые данным лечением не улучшаются. При этом могут динамически предоставляться возможные объяснения наблюдаемых аномалий. В будущем компьютеры должны приобрести способность формировать точные и содержательные документы, подобные выписным эпикризам, составляемым опытными больничными врачами.

6.3. Интерфейс и особенности электронной истории болезни

Медицинская карта включает в себя информацию, возникающую при амбулаторной и стационарной работе с пациентом. Естественная взаимосвязь элементов медицинской карты позволяет реализовать не только традиционный механизм передачи информации посредством выписок из амбулаторной карты и истории болезни, но и более эффективный и полезный полный доступ к любой информации о пациенте в пределах полномочий пользователя системы.

Разработанная авторами монографии электронная медицинская карта [93] включает в себя информацию, получаемую при амбулаторной и стационарной работе с пациентом (рис.6.1). Естественная взаимосвязь

элементов медицинской карты позволяет реализовать традиционный механизм передачи информации посредством выписок из амбулаторной карты и истории болезни. Предусмотрен эффективный полный доступ к любой информации о пациенте в пределах полномочий пользователя системы. При этом лечащий врач, работающий с записями амбулаторной карты пациента, имеет возможность анализировать не только агрегированную информацию выписки из истории болезни, но и всю первичную информацию, содержащуюся в любой истории болезни, включенной в медицинскую карту.

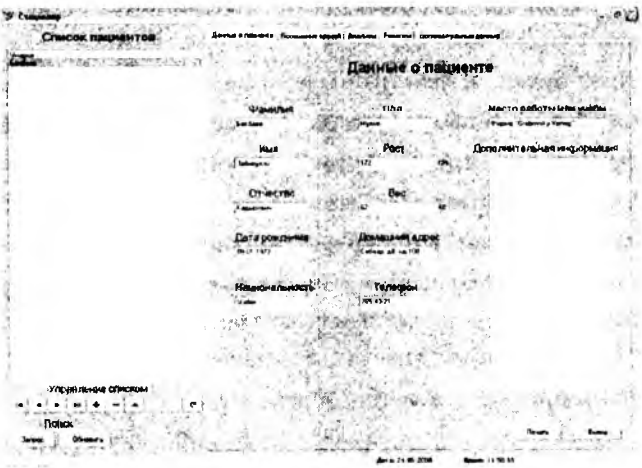


Рис.6.1. Информация, получаемая при амбулаторной и стационарной работе с пациентом

Медицинская карта представляет собой пополняемый набор конфигурируемых групп, содержащих важные, с точки зрения лечебно-диагностического процесса, элементы медицинской карты (рис.6.2).

Рис.6.2. Элемент медицинской карты – «Анализ крови»

Стандартная конфигурация включает в себя следующие группы: физиологические признаки параметры, клинические признаки заболевания, лабораторные показатели, данные инструментальных исследований, проводимое лечение. Каждый врач может произвольным образом, в соответствии со своими представлениями о значимости отслеживаемых параметров, результатов исследований и пр., наполнять группу элементами медкарты в зависимости от конкретного случая – нозологической формы, особенностей лечения заболевания и т.д. (рис.6.3).

Рис.6.3. Диагноз нозологической формы

Предусмотрена возможность компоновки визуального отображения для удобства отслеживания критических событий и тенденций лечебного процесса – например, поместить рядом физиологические параметры и схему исполнения медикаментозных назначений. Выделяются случаи отличия от нормы, отказов и задержек при проведении исследований и исполнении назначений и т.п.

Врач, имея такой достаточно мощный и гибкий инструмент отображения содержимого медицинского технологического процесса, получает наглядную возможность отслеживать значимые и критические изменения параметров организма, оценивать степень воздействия применяемого лечения, оперативно реагировать на изменения при исполнении лечебных назначений, получать и сопоставлять информацию о пациенте из различных источников – автоматизированных рабочих мест врачей, медсестёр, биохимических лабораторий и т.д. Врач имеет возможность наблюдать в динамике степень соответствия течения заболевания выведенной им диагностической формуле и получает достаточную информационную поддержку для внесения корректив в клиническую картину пациента.

Заключение

Медицина всегда активно впитывала достижения физики, химии, математики. Так создавались рентгенотерапия и рентгенодиагностика, медицинская радиология, лазерная медицина. Не стали исключением и информационные технологии.

Поэтому в монографии нашли отражение результаты научных исследований по некоторым основным направлениям применения информационных технологий в здравоохранении: автоматизация обработки документации в лечебно-профилактических учреждениях на примере электронной истории болезни; экспертные системы и программные комплексы для лечения и диагностики некоторых заболеваний; статистические системы для анализа результатов медико-биологических исследований ферментной диагностики сердечно-сосудистой системы; нейросетевые технологии для анализа и обработки биомедицинских данных.

Представленные авторами монографии результаты научных исследований внедрены в медицинскую практику для решения задач дифференциальной диагностики, выбора оптимального режима лечения и прогнозирования исхода отдельных заболеваний.

Литература

1. Математико-статистические методы в клинической практике/Под ред. В.И. Кувакина. –С.-Пб.: ВМедА,1993. –199 с.

2. Юнкеров В.И., Григорьев С.Г. Математико-статистическая обработка данных медицинских исследований. –С.-Пб.: ВМедА, 2002. – 266 с.

3. Петри А., Сэбин К. Наглядная статистика в медицине. –М.: Издательский дом ГЭОТАР-МЕД, 2003. –142 с.

4. Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. –Киев: МОРИОН, 2001. –408 с.

5. Платонов А.Е. Статистический анализ в медицине и биологии: задача, терминология, логика, компьютерные методы.–М.: РАМН, 2000. – 52 с.

6. Поллок М.Л., Шмидт Д.Х. Заболевания сердца и реабилитация. – Киев. Олимпийская литература, 2000. –408 с.

7. Магрупов Т.М., Васильева С.А., Леонова Г.Ю. Ферментная диагностика инфаркта миокарда с использованием статистических методов//Республ. межвуз. сб. науч. тр. «Актуальные вопросы в области технических и социально-экономических наук». – Ташкент, 2010. – С.65-67.

8. Реброва О.Ю. Статистический анализ медицинских данных. Применение пакета прикладных программ Statistica. –М.: Медиа Сфера, 2002. –312 с.

9. Галушкин А.И. Теория нейронных сетей. Кн. 1: Учеб. пособие для вузов. – М.: ИПРЖР, 2001. – 385 с.

10. Головко В.А. Нейронные сети: обучение, организация и применение. Кн.4:Учеб. пособие для вузов. –М.: ИПРЖР, 2001. –256 с.

11. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. –М.: Горячая линия - Телеком, 2001. –382 с.

12. Каллан Р. Основные концепции нейронных сетей. –М.: Вильямс, 2001. –288 с.

13. Миркес Е.М. Нейрокомпьютер. Проект стандарта. –Новосибирск: Наука, 1999. –337 с.

14. Вороновский Г.К., Махотило К.В. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности. – Харьков: ОСНОВА, 1997. –112 с.

15. Терехов В.А., Ефимов Д.В., Тюкин И.Ю. Нейросетевые системы управления. –М.: Высшая школа, 2002. –184 с.

16. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. – М.: Мир, 1992. –240 с.

17. Хайкин С. Нейронные сети: полный курс. –М.: Вильямс,2006.– 1104 с.

18. Хетагурова А.К., Слепушенко И.О. и др. Применение нейроинформационных технологий в паллиативной медицине//Сестринское дело. –2003. –№6. –С. 7-9.

19. Нейронные сети (http://www.codenet.ru/progr/alg/ai/).

20. Нейросети и проблема искусственного интеллекта (http://elik.ru/papers/nn_ai/).

21. Современные направления развития нейрокомпьютерных технологий в России (http://neurnews.iu4.bmstu.ru/neurnews.html).

22. Нейрокомпьютеры (http://neurnews.iu4.bmstu.ru).

23. Фролов Ю.В. Интеллектуальные системы и управленческие решения. - М.: МГПУ, 2000. –294 с.

24. Дмитриенко В.Д. Основы нейронных сетей.– Белгород, 2001.– 159 с.

25. Гайдышев И. Анализ и обработка данных: специальный справочник. –С.-Пб.: Питер, 2001.–752 с.

26. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. –Новосибирск: Наука, 1996. 276 с.

27. Ежов А., Чечеткин В. Нейронные сети в медицине//Открытые системы. – 1997. – №4. – С.34-37.

28. Cagnoni S., Coppini G., Rucci M. et al. Neural network segmentation of magnetic resonance spin echo images of the brain//J. Biomed. Eng.- 1993.- V.15, №.5.- P.355-362.

29. Astion M.L., Wener M.H., Thomas R.G. Application of neural networks to the classification of giant cell arteritis//Arthritis Reum.- 1994.- V.37, №.5.- P.760-770.

30. Baxt W.G. A neural network trained to identify the presence of myocardial infarction bases some decisions on clinical associations that differ from accepted clinical teaching//Med. Decis. Making.- 1994.- V.14, №3.- P.217-222.

31. Baxt W.G. Complexity, chaos and human physiology: the justification for non-linear neural computational analysis //Cancer Lett.- 1994.- V.77, №2-3.- P.85-93.

32. Baxt W.G. Use of an artificial neural network for the diagnosis of myocardial infarction//Ann. Intern. Med.- 1991.- V.115, №11.- P.843-848.

33. Guo Z., Durand L.G., Lee H.C. et al. Artificial neural networks in computer-assisted classification of heart sounds in patients with porcine bioprosthetic valves//Med. Biol. Eng. Comput.- 1994.- V.32, №3.- P.311-316.

34. Barschdorff D., Ester S., Dorsel T et al. Phonographic diagnostic aid in heart defects using neural networks//Biomed. Tech. Berlin.- 1990.- V.35, №11.- P.271-279.

35.Okamoto Y., Nakano H., Yoshikawa M. et al. Study on decision support system for the interpretation of laboratory data by an artificial neural network// Rinsho. Byori.- 1994.- V.42, №2.- P.195-199.

36. Maclin P.S., Dempsey J. Using an artificial neural network to diagnose hepatic masses//J. Med. Syst.- 1992.- V.16, №5.- P.215-225.

37. Rinast E., Linder R., Weiss H.D. Neural network approach for computer-assisted interpretation of ultrasound images of the gallbladder//Eur. J. Radiol.- 1993.- V.17, №3.- P.175-178.

38. Modai I., Stoler M., Inbar-Saban N. et al. Clinical decisions for psychiatric inpatients and their evaluation by a trained neural network//Methods Inf. Med.- 1993.- V.32, №5.- P.396-399.

39. Berrios G.E., Chen E.Y. Recognising psychiatric symptoms. Relevance to the diagnostic process//Br. J. Psychiatry.- 1993.- V.163.- P.308-314.

40. Mulsant B.H. A neural network as an approach to clinical diagnosis//MD Comput.- 1990.- V.7, №1.- P.25-36.

41. Ercal F., Chawla A., Stoeker W.V. et al. Neural network diagnosis of malignant melanoma from color images//IEEE Trans. Biomed. Eng.- 1994.- V.41, №9.- P.837-845.

42. Lee H.-L., Suzuki S., Adachi Y. et al. Fuzzy Theory in Traditional Chinese Pulse Diagnosis//Proceedings of 1993 International Joint Conference on Neural Networks, Nagoya, 1993.- V.1.- P.774-777.

43. Yang T.-F., Devine B., Macfarlane P.W. Combination of artificial neural networks and deterministic logic in the electrocardiogram diagnosis of inferior myocardial infarction//Eur. Heart J.- 1994.- V.15.- Abstr. Supplement XII-th World Congress Cardiology (2408).- P.449.

44. Yang T.-F., Devine B., Macfarlane P.W. Computer assisted electrocardiogram detection of myocardial infarction by software based artificial neural networks//Eur. Heart J.- 1994.- V.15.- Abstr. Supplement XII-th World Congress Cardiology (2407).- P.449.

45.Yang T.-F., Devine B., Macfarlane P.W. Deterministic logic versus software-based artificial neural networks in the diagnosis of atrial fibrillation//J. Electrocardiol.- 1993.- V.26, Suppl.- P.90-94.

46.Yang T.-F., Devine B., Macfarlane P.W. Differentiation of atrial fibrillation from sinus rhythm and (supra)ventricular extrasystoles using software based neural networks//Eur. Heart J. Abstr. Supplement XVth Congress of the European Society of Cardiology.- 1993, V.14.- P.235.

47. Hoher M., Kestler H.A., Palm G. et al. Neural network based QRS classification of the signal averaged electrocardiogram//Eur. Heart J. Abstr. Supplement XII-th World Congress Cardiology.- 1994, V.15.- P.114.

48. Nakajima H., Anbe J., Egoh Y. et al. Evaluation of neural network rate regulation system in dual activity sensor rate adaptive pacer//European Journal of Cardiac Pacing and Electrophysiology.- Abstracts of 9th International Congress, Nice Acropolis - French, Rivera, 1994.- P.54.

49. Gross G.W., Boone J.M., Greco-Hunt V. et al. Neural networks in radiologic diagnosis. II. Interpretation of neonatal chest radiographs//Invest. Radiol.- 1990.- V.25, №9.- P.1017-1023.

50. Floyd C.E.Jr., Lo J.Y., Yun A.J. et al. Prediction of breast cancer malignancy using an artificial neural network//Cancer.- 1994.- V.74, №11.- P.2944-2948.

51. Fujita H., Horita K., Endo T. et al. Neural Network Approach to Classification of Benign and Malignant Tumours in Mammograms//Med. Imag. Tech.- 1992.- V.10.- P.126-129.

52. Reinbnerger G., Weiss G., Werner-Felmayer G. et al. Neural networks as a tool for utilizing laboratory information: comparison with linear discriminant analysis and with classification and regression trees//Proc. Natl. Acad. Sci., USA.- 1991.- V.88, №24.- P.11426-11430.

53. Molnar B., Szentirmay Z., Bodo M. et al. Application of multivariate, fuzzy set and neural network analysis in quantitative cytological examinations// Anal. Cell Pathol.- 1993.- V.5, №3.- P.161-175.

54. Vertosick F.T., Rehn T. Predicting behavior of an enzyme-linked immunoassay model by using commercially available neural network software// Clin. Chem.- 1993.- V.39, №12.- P.2478-2482.

55. Шварц Э., Трис Д. Программы, умеющие думать//Бизнес Уик.– 1992.– №6.– С.15-18.

56. http://www.statsoft.ru/home/textbook/default.htm.

57. Боровиков В. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. –С.-Пб.: Питер, 2001.–656 с.

58. Нейронные сети: STATISTICA Neural Networks. –М.: Горячая линия – Телеком, 2001.–92 с.

59. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. –М.: МедиаСфера, 1998. –352 с.

60. Васильева С.А. Построение нелинейных моделей сердечно-сосудистых заболеваний в пакете STATISTICA//Вестник ТашГТУ.– 2011. – №3-4. – С.25-29.

61. Тихонов Э.Е. Методы прогнозирования в условиях рынка: Учеб. пособие. - Невинномысск, 2006. – 221 с.

62. Потемкин В. Г. Система MATLAB: Справ. пособие. –М.: ДИАЛОГ-МИФИ, 1997. –350 с.

63. Гультяев А. Визуальное моделирование в среде MATLAB: Учебный курс. – С.-Пб.: Питер, 2000. –432 с.

64. Потемкин В. Г. Система инженерных и научных расчетов MATLAB 5.x: В 2-х т. Т.1. –М.: Диалог-МИФИ, 1999. –366 с.

65. http://www.matlab.exponenta.ru

66. Мурашко В. В., Струтынский А. В. Электрокардиография. –М.: Медицина, 1987.– 256 с.

67. Тигомир Л. И., Кнешпо П. Математическое моделирование биоэлектрического генератора сердца.– М.: Наука, Физматлит,1999.– 448 с.

68. Матрунов Т.М., Васильева С.А., Гаибназаров Б.Б, Маурус Фрей Синтез электрокардиосигналов на основе линейно-квадратичной интерполяции в системе MATLAB 6.0//Государственный стандарт Узбекистана. Ташкент. – 2007. – №1.– С.18-19.

69. Стечкин С.Б., Субботин Ю.Н. Сплайны в вычислительной математике. –М.: Наука. 1976. –248 с.

70. Матрунов Т.М., Васильева С.А., Соловьёв М.Ю. Алгоритм синтеза электрокардиосигналов с использованием кубических сплайнов //Республ. межвуз.сб.науч. тр. «Актуальные вопросы в области технических и социально-экономических наук», Ташкент, 2008. – С.108-112.

71. Гребенев А. Л. Пропедевтика внутренних болезней: Учебник. –М.: Медицина, 2001.–592 с.

72. Мухин Н.А., Моисеев В.С. Пропедевтика внутренних болезней: Учебник. –М.: ГЭОТАР-Медиа, 2004. –768 с.

73. Магрупов Т.М., Васильева С.А. Программное обеспечение для радионуклидного исследования функционального состояния печени с портальной гипертензией//Межд.сб.науч.тр.«Медицинские приборы и технологии». –Тула, Тульский государственный университет, 2011. – С.187-191.

74. Радионуклидное исследование функционального состояния печени с портальной гипертензией. Свидетельство об официальной регистрации программ для ЭВМ №DGU 02149, 2011/Магрупов Т.М., Васильева С.А.

75. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование. – М.: Вильямс, 2006. - 152 с.

76. Смолин Д.В. Введение в искусственный интеллект. М.: Физматлит, 2004.–208 с.

77. Назаренко Г.И., Осипов Г.С. Медицинские информационные системы и искусственный интеллект. Вып. 3: Науч. пособ. - М.: Медицина XXI, 2003. –240 с.

78. Овчинников Ю.М. Отоларингология. – М.: Медицина, 1995. -232 с.

79. Джафек Б., Старк Э. Секреты оториноларингологии. – М.: Бином, 2001. –256 с.

80. Экспертная система оториноларингологических заболеваний. Свидетельство об официальной регистрации программ для ЭВМ №DGU 02058, 2010/Магрупов Т.М., Васильева С.А., Голубев А.С.

81. Светухин А.М., Звягин А.А., Слепнев С.Ю. Системы объективной оценки тяжести состояния больных//Хирургия. – 2002. – №9. – С. 50-57.

82. Корячкина В.А., Страшнова В.И., Чуфарова В.П. Клинические функциональные и лабораторные тесты в анестезиологии интенсивной терапии. – С.-Пб.: Санкт-Петерб. мед. изд-во, 2001. – 326 с.

83. Долина О.А. Анестезиология и реаниматология. –М., 1998.– 544 с.

84. Комаров Ф.И., Коровкин Б.Ф. Биохимические показатели в клинике внутренних болезней: Справочник. –М.: МЕДпресс, 2000. – 232 с.

85. Полушин Ю.С. Руководство по анестезиологии и реаниматологии. С.-Пб. 2004. – 919 с.

86. Рафф Г. Секреты физиологии. –М.; С.-Пб.: БИНОМ–Невский Диалект.– 2001. – С. 448.

87. Генес В.С. Некоторые простые методы кибернетической обработки данных физиологических и диагностических исследований. – М.: Наука, 1967. – 208 с.

88. Магрупов Т.М., Васильева С.А. Абдазимов А.А. Программный комплекс автоматизированной обработки данных клинических функциональных исследований дыхания//Республ.межвуз.сб.науч.тр., «Актуальные вопросы в области технических и социально-экономических наук». Ташкент, 2006. –С. 90-93.

89. Томсон Л., Веллинг Л. Разработка WEB-приложений на PHP и MySQL. –С.-Пб., ДиаСофЮП, 2003.–672 с.

90. Назаренко Г.И., Гулиев Я.И., Ермаков Д.Е. Медицинские информационные системы: теория и практика.. –М.:Физматлит,2005.–320 с.

91. Рог Г. З. Медицинские информационные системы: Учеб. особие. Новосибирск: Изд-во НГТУ, 2005. – 70 с.

92. Тавровский В.М. Лечебно-диагностический процесс: Теория. Алгоритмы. Автоматизация. –Тюмень: СофтДизайн, 1997. – 320 с.

93. Магрупов Т.М., Васильева С.А., Закиров К.Н. Универсальный интерфейс рабочего места врача//Республ.межвузов.сб.науч.тр. «Актуальные вопросы в области технических и социально-экономических наук». Ташкент, 2008. –С. 102-104.

Содержание

Введение. 3

Глава 1. Многофакторный корреляционный и регрессионный анализ данных медицинских исследований 8

1.1. Задачи исследований сложных систем. 8

1.2. Ферментная диагностика инфаркта миокарда с использованием статистических методов. 15

Глава 2. Нейросетевые технологии для анализа и обработки медицинских данных. 21

2.1. Искусственные нейронные сети: строение, принципы работы и обучение. 22

2.2. Нейросетевые подходы обработки медицинской информации. . 28

2.3. Диагностические медицинские программы, основанные на нейроприложениях. 32

2.4. Информационные технологии Data Mining. 40

2.5. Диагностика сердечной патологии с помощью STATISTICA Data Miner. 43

2.6. STATISTICA Neural Networks в нейросетевых вычислениях. . . 46

2.7. Задачи регрессии в пакете STATISTICA Neural Networks. 51

2.8. Аппроксимация функций в пакете STATISTICA 58

2.9. Модели сердечно-сосудистых заболеваний в пакете STATISTICA.. 61

Глава 3. Методы обработки медицинских изображений и распознавания образов в системе MATLAB 68

3.1. Среда системы MATLAB. 69

3.2. Обработка биомедицинских изображений с использованием системы MATLAB. 77

3.3. Синтез электрокардиосигналов в системе MATLAB. 79

3.4. Анализ сцинтиграмм печени с использованием среды MATLAB 86

Глава 4. Экспертные системы в медицине. 93
4.1. Экспертные системы – эффективное диагностическое средство в медицине. 94
4.2. Алгоритмическое и программное обеспечение экспертной системы диагностики оториноларингологических заболеваний. 103
Глава 5. Способы анализа и обработки медико-биологических данных. 109
5.1. Системы оценки состояния больного. 109
5.2. Алгоритмическое и программное обеспечение оценки тяжести состояния пациентов. 115
5.3. Автоматизация обработки данных при оценке кислотно-щелочного состояния организма. 118
5.4. Автоматизированная обработка данных клинических функциональных исследований дыхания. 121
Глава 6. Автоматизированные системы ведения истории болезни. . . 125
6.1. Основные факторы ведения истории болезни. 126
6.2. Фундаментальные вопросы разработки и внедрения автоматизированных систем ведения истории болезни. 130
6.3. Интерфейс и особенности электронной истории болезни. 137
Заключение. 141
Литература. 142

Анализ и обработка

медико-биологической информации

Магрупов Талат Мадиевич, Васильева Стелла Александровна,

Магрупова Малохат Талатовна

Редактор: Н.С. Покачалова

Технический редактор: Косимов Б.

Подписано к печати 15.08.2012 г. Формат 60x84 1/16.
Объем 8,85 п.л. Тираж 50 экз. Заказ № 500.

Отпечатано в типографии ТГТУ. г.Ташкент,
ул.Талабалар 54. тел: 246-63-84.