

Костюкова Нина Ивановна, Кудинов Андрей Евгеньевич

**АВТОМАТИЗАЦИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ В ОБЛАСТИ МЕДИЦИНЫ С ПРИМЕНЕНИЕМ
ТЕХНОЛОГИИ DATA MINING**

Адрес статьи: www.gramota.net/materials/1/2010/3-1/7.html

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

Альманах современной науки и образования

Тамбов: Грамота, 2010. № 3 (34): в 2-х ч. Ч. I. С. 22-24. ISSN 1993-5552.

Адрес журнала: www.gramota.net/editions/1.html

Содержание данного номера журнала: www.gramota.net/materials/1/2010/3-1/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: almanac@gramota.net

УДК 65.011.5

*Нина Ивановна Костюкова, Андрей Евгеньевич Кудинов**Институт вычислительной математики и математической геофизики СО РАН*

АВТОМАТИЗАЦИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ В ОБЛАСТИ МЕДИЦИНЫ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИИ *DATA MINING*[©]

Описание предметной области

В процессе установления диагноза и выработки рекомендуемого лечения врач часто сталкивается с целой последовательностью сложных решений. В медицинской литературе находят свое отражение и крайне трудные случаи, в которых врачу часто приходится принимать ответственные решения. Отсюда с очевидностью вытекает потребность аналитически сформулировать такого рода задачи, с которыми сталкивается врач. Цель нашей работы - сформулировать в математических терминах логические основы диагностики и выбора метода лечения, используя, как можно шире существующие методы математического моделирования и реализовать их на ЭВМ в рамках современных информационных технологий. Мы разработали архитектуру открытой автоматизированной системы. По нашему мнению эта система будет допускать практическое применение в двух направлениях. Первое связано с обучением студентов-медиков, второе относится к применению вычислительных устройств в качестве подсобного средства в руках врача. Система реализуется на платформе .Net на языке программирования C#. Ее архитектура: функциональное ядро, содержащее программную реализацию методов диагностики и выбора методов лечения, пользовательский интерфейс, база данных и база знаний предметной области, криптографический модуль защиты информации. Система открыта и допускает расширение на любом уровне.

Принципы выбора математических моделей

Опишем несколько проблем, над которыми мы уже работали и получили некоторые результаты.

Изучение сигналов и их распознавание тесно связано с проблемами поведения центральной нервной системы и мозга, памяти и обучения и самих процессов принятия решений. Попытки понять или хотя бы смоделировать эти явления привели к изучению автоматов, машин, способных принимать решения, и, конечно, к абстрактному исследованию постановки и решения задач.

Переходя от изучения макропроцессов к изучению микропроцессов, мы коснемся вопроса размножения клеток. Как объяснить ту невероятную точность, с которой происходит этот процесс из поколения в поколение? Для исследования рассматриваемых вопросов используется сложный алгебраический и топологический аппарат. При изучении передачи информации от родителей к потомству мы сталкиваемся с рядом фундаментальных проблем связи и управления.

Как реагируют живые системы на химические воздействия и облучения? Здесь одна из главных задач - усовершенствование методов лечения рака с помощью радиационной терапии и химиотерапии. Определить, до какой степени управляемы ситуации (мы используем метод линейного программирования).

Использование открытой системы ни в коей мере не означает передачу ей функций врача. Наоборот, роль врача становится при этом еще более трудной. Врачу придется быть более образованным: помимо своих профессиональных знаний, ему придется овладеть и теми новыми методами, которые будут заложены в этой системе. Однако мы надеемся, что ценой дополнительных усилий можно будет добиться выигрыша в точности диагноза и эффективности лечения.

Логика в диагностике

Для решения задачи установления диагноза используются следующие источники: медицинские знания, симптомы, обнаруживаемые у больного. Медицинские знания содержат некоторую информацию о соотношениях между симптомами и болезнями. Симптомы дают дальнейшую информацию о самом больном. Исходя из этих двух источников информации, методами исчисления предложений ставится диагноз. Пусть n - число симптомов, а m - число болезней. Обозначим симптомы $S(1), S(2), \dots, S(n)$, а болезни - $D(1), D(2), \dots, D(m)$.

По отношению к определенному больному символ $S(i)$ будет означать, что этот больной обнаруживает i -й симптом, а символ $D(j)$ - что больной страдает j -м из перечисленных заболеваний. Представим себе всевозможные комбинации перечисленных симптомов и заболеваний. Которые могут быть обнаружены у данного пациента; общее число таких комбинаций равно 2^{n+m} . Любую такую комбинацию мы условимся называть комплексом симптомов и заболеваний и обозначать C_j^i ; произвольную комбинацию симптомов будем называть комплексом симптомов и обозначать s^i , а комбинацию заболеваний - комплексом заболеваний d_j . В этих обозначениях $s^i \cdot d_j = C_j^i$.

Составим матрицу из нулей и единиц. Столбцы этой матрицы изображают все мыслимые комплексы симптомов и заболеваний. Назовем ее логическим базисом данных симптомов и заболеваний. Мы употребили слово «мыслимые», а не «возможные», так как не все они могут иметь место в действительности. Указать какие комплексом симптомов и заболеваний могут иметь место, какие нет, - дело медицины. Предположим, что медицина дает следующие указания:

1. У пациента с заболеванием $D(1)$ и одновременно $D(2)$ или $D(3)$ должны обнаруживаться симптомы $S(1), S(3)$.

2. В отсутствие $D(2)$ не может обнаруживаться симптом $S(2)$.

3. У пациента с заболеваниями $D(2), D(3)$ при отсутствии $D(1)$ должен обнаруживаться симптом $S(3)$.

4. У пациента обнаруживается хотя бы один из симптомов и есть хотя бы одно заболевание.

Мы понимаем, конечно, что наличие симптома может свидетельствовать о наличии заболевания, но отсутствие явных симптомов не означает отсутствия заболевания.

Переведем указания врача, пункты 1, 2, 3, 4, в математическую форму записи. В системе эту функцию выполняет «интеллектуальный блок».

1. $D(1) \cdot [D(2) + D(3)] \rightarrow S(1) \cdot S(3)$.

2. $\overline{D(2)} \rightarrow \overline{S(2)}$.

3. $\overline{D(1)} \cdot D(2) \cdot D(3) \rightarrow S(3)$.

4. $[S(1) + S(2) + S(3)] \cdot [D(1) + D(2) + D(3)]$.

Логическое произведение этих предложений является булевой функцией.

$E = E(S(1), \dots, S(n), D(1), \dots, D(m))$, выражающей соотношения между заболеваниями и симптомами, почерпнутые из медицинской науки.

Пусть теперь перед нами конкретный больной, и мы должны поставить диагноз. Какие-то симптомы у него обнаружены, какие-то другие - нет; в общем случае есть еще и третья группа симптомов, относительно которых в данный момент мы не располагаем сведениями. Такая совокупность симптомов, быть может, не полная, называется профилем симптомов больного и обозначается G . Одновременно рассматривается совокупность заболеваний (быть может, неполная) - профиль заболеваний f . Оба профиля вместе образуют профиль симптомов и заболеваний $f \cdot G$. Профиль G есть булевская функция симптомов.

Логическое произведение этих предложений является булевой функцией.

$E = E(S(1), \dots, S(n), D(1), \dots, D(m))$,

выражающей соотношения между заболеваниями и симптомами, почерпнутые из медицинской науки.

Пусть теперь перед нами конкретный больной, и мы должны поставить диагноз. Какие-то симптомы у него обнаружены, какие-то другие - нет; в общем случае есть еще и третья группа симптомов, относительно которых в данный момент мы не располагаем сведениями. Такая совокупность симптомов, быть может, не полная, называется профилем симптомов больного и обозначается G . Одновременно рассматривается совокупность заболеваний (быть может, неполная) - профиль заболеваний f . Оба профиля вместе образуют профиль симптомов и заболеваний $f \cdot G$. Профиль G есть булевская функция симптомов

$G = G(S(1), \dots, S(n))$.

Теперь мы можем сформулировать логическую задачу: по заданному профилю симптомов G и функции E найти профиль заболеваний f как булевскую функцию аргументов $D(1), \dots, D(m)$;

$f = f(D(1), \dots, D(m))$.

Медицинские проблемы формулируются врачом, на входном языке, разработанным к базе знаний. Затем при помощи логических выводов эти данные переводятся в логические записи, на базе которых строятся логические базисы заболеваний, которые используются в диагностики заболеваний.

Будет реализован и отлажен в промышленности алгоритм лечения рака. Данный алгоритм строится на методах линейного программирования.

Нами разработаны и реализованы несколько алгоритмов функционального ядра. К данному моменту разработан алгоритм лечения рака (используется Симплекс-метод). Разработан алгоритм выбора лечения в условиях неизвестности. Получены уравнения математического ожидания излеченных больных. Алгоритм построен с использованием теории игр. Разработан алгоритм выбора лечения в условиях риска. Получено уравнение математического ожидания процента больных, излеченным данным методом. Используется теория вероятностей.

Учет индивидуальных особенностей больного

Состояние здоровья определяется физиологическими и биохимическими показателями. Изменение или стандартное отклонение такого показателя у отдельного индивидуума меньше стандартного отклонения этого показателя у всей популяции, и среднее значение такого показателя у отдельного индивидуума может не совпадать со средним, вычисленным для всей популяции. Поэтому значение, лежащее в пределах нормы для популяции, может не быть нормальным для отдельного индивидуума; возможно и обратное. Вообще говоря, такого рода показатели варьируют от индивидуума к индивидууму гораздо больше, чем принято считать.

3. Уилльямс [3, р. 227] замечает, что практически любой человеческий организм представляет собой в том или ином смысле отклонение от нормы. Поэтому, производя какое-либо исследование, мы должны исходить из норм для данного индивидуума, а не из норм, установленных для популяции, как делают обычно. Изучение биохимической индивидуальности стало возможно лишь в последнее время с появлением таких методов исследования, как хроматография, изотопные методы и физические методы анализа и разделения.

Другая причина, заставляющая нас ориентироваться на индивидуальные нормы, состоит в том, что биохимические и физиологические показатели изменяются с возрастом. Вместе с ними меняется и индивидуальная восприимчивость к той или иной болезни. Это обстоятельство ясно выступает при изучении данным о смертности: именно индивидуальные отклонения в процессе «выхода из строя», или дегенерации, различных биохимических и физиологических систем объясняют «квантовый» характер дифференциальных кривых смертности [2, р. 129]. Обнаружив на ранней стадии этого процесса значительное отклонение какого-либо биохимического показателя от нормы, присущей данному индивидууму, можно прибегнуть к предупредительным мерам. Но в настоящее время имеется мало данных о биохимических и физиологических состояниях в начальных и последующих стадиях заболеваний.

Авторы считают, что в данной работе новыми являются следующие положения и результаты: Изучена современная информационная технология. В долгосрочной перспективе будущее *Data Mining* является действительно захватывающим - это может быть поиск интеллектуальными агентами, как новых видов лечения различных заболеваний, так и нового понимания природы вселенной. Мы же применили *Data Mining* для автоматизации медицинской диагностики и выбора методов лечения. Результаты докладывались в ООО «Вектор», Научноград Кольцово Новосибирской области и на Седьмой международной конференции памяти академика А. Е. Ершова «Перспективы систем информатики» [1, с. 72-76].

Список литературы

1. **Костюкова Н. И.** Система принятия решений по технологии Data Mining // Перспективы систем информатики: материалы Седьмой международной конференции памяти академика А. Е. Ершова. Новосибирск, 2009. С. 72-76.
2. **Sacher G. A.** Reparable and irreparable injury // Radiation biology and medicine. 1999. P. 297.
3. **Williams R. J.** Biochemical individuality. New York, 1999. P. 129.

УДК 536.2

*Игорь Евгеньевич Лобанов, Леонид Михайлович Штейн
Московский авиационный институт (государственный технический университет)
Московская академия коммунального хозяйства и строительства*

ТЕОРИЯ ИНТЕНСИФИЦИРОВАННОГО ТЕПЛООБМЕНА И ЭФФЕКТИВНОСТИ ЕГО ПРИМЕНЕНИЯ ДЛЯ ПЕРСПЕКТИВНЫХ КОМПАКТНЫХ ТЕПЛООБМЕННЫХ АППАРАТОВ, ПРИМЕНЯЕМЫХ В СОВРЕМЕННОМ МЕТАЛЛУРГИЧЕСКОМ ПРОИЗВОДСТВЕ[©]

Данное исследование выполнено при поддержке Российского фонда фундаментальных исследований (грант РФФИ № 09-08-00440), Президента РФ по государственной поддержке научных исследований молодых российских учёных - докторов наук (грант МД № 1420.2008.8)

1. Постановка задачи

Актуальность проблемы интенсификации теплообмена в теплообменниках, используемых в современном металлургическом производстве, наиболее подробно была показана в монографиях [11; 12], где указывалось на преимуществе теоретического исследования интенсифицированного теплообмена над экспериментальным.

Теоретическое исследование интенсифицированного теплообмена в теплообменниках для металлургического производства основывается на имеющихся разработках в этой области [4; 5; 6; 7; 10; 23], которые успешно продолжают в настоящее время при поддержке гранта РФФИ № 09-08-00440 и гранта Президента РФ № МД 1420.2008.8.

Постановка теоретической задачи заключается в математическом моделировании интенсифицированного теплообмена в каналах теплообменных аппаратов, применяемых в современном металлургическом производстве в целях улучшения эффективности и снижения массогабаритных показателей. Далее следует рассмотреть теплообменники, теплообмен в которых необходимо интенсифицировать.