

УДК 616-71 ГРНТИ 76.13.15

Забезжайло М.И., Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Вавилова, д.40;
m.zabezhailo@yandex.ru

Трунин Ю.Ю., Национальный медицинский исследовательский центр нейрохирургии имени академика Н.Н.Бурденко Минздрава РФ, Москва, 4-я Тверская-Ямская, д.16; ytrunin@nsi.ru

О доказательности медицинского диагноза: интеллектуальный анализ эмпирических данных о пациентах в выборках ограниченного размера

Аннотация

Обсуждаются возможности расширения используемых в доказательной медицине аргументационных возможностей применением логико-математических средств обоснования утверждений в открытых теориях. Представлена схема развиваемого на этой основе интеллектуального анализа эмпирических данных для задач медицинской диагностики. Демонстрируется эффективность и практическая значимость развиваемого подхода на примере проблемы диагностики псевдопрогрессии опухолей головного мозга.

Ключевые слова:

интеллектуальный анализ данных, доказательная медицина, системы поддержки принятия врачебных решений, псевдопрогрессия опухолей головного мозга

Доказательная медицина (EBM) – очевидный шаг вперед в развитии медицинских технологий, позволяющий существенным образом повысить качество анализа данных в процессе принятия врачебных решений. Исторически сложилось так, что в части математических оснований доминирующим EBM-инструментарием обработки эмпирических данных оказались статистические методы. Однако, в целом ряде практически значимых случаев (например, в задачах высокотехнологичной диагностики и лечения патологий нейрохирургического характера) приходится иметь дело с характеризующими состояние пациентов выборками эмпирических данных, имеющими ограниченный (и при этом – достаточно медленно пополняемый описаниями новых прецедентов) размер. Статистические методы анализа таких коллекций данных, к сожалению, не позволяют формиро-

вать неоспоримые диагностические заключения персонализированного характера. Наиболее трудные (в части неоспоримости предлагаемых решений) вопросы здесь оказываются связанными с

- *точностью* формируемого в результате анализа данных (АД) именно такими методами заключения (например, диагноза в случае конкретного пациента в его конкретном состоянии), характеризующего *вероятностью* ошибки,
- *персонализацией* диагностического заключения (ориентированностью выносимого заключения именно на конкретного пациента в его конкретном состоянии, а не только с фокусировкой на статистически значимую группу лиц, к которой может быть отнесен данный пациент),

- *неформальной интерпретируемостью (объясняемостью в содержательных медицинских терминах, понятиях и категориях, а не в терминологии процедурно-алгоритмической схемы¹, используемой для математического анализа данных) результатов АД, лежащих в основе соответствующего диагностического заключения.*

Возможности обойти подобные трудности предоставляет представление о доказуемости (утверждения) как о (его) *неоспариваемости* на имеющихся данных. Так в случае работы с открытыми (допускающими пополнение новыми данными) *предметными областями (ПО)*, - а медицина представляет собою один из наиболее хорошо известных примеров ПО такого типа, - для формализованного описания накапливаемых эмпирических данных естественно использовать так называемые *открытые теории (ОТ)*. Каждая ОТ – это открытое (пополняемое при появлении новых данных) множество таких истинных (на имеющихся данных) формул, что любой из зафиксированных на текущий момент в рассматриваемой предметной области фактов может быть представлен как *логическое следствие* этого множества формул. Процедурная схема, позволяющая накапливать коллекции эмпирических фактов (баз фактов - БФ), рассматривать их пополнения новыми фактами, выделять в них эмпирические зависимости (истинные в текущей коллекции фактов формулы - ЭЗ), собирать их в такие семейства, что любой факт из текущей БФ является логическим следствием ЭЗ из такого семейства (открытой теории, описывающей текущее состояние БФ) в общем виде характеризует предлагаемую нами схему интеллектуального анализа данных (ИАД), ориентированную на решение задач диагностического характера.

Используемая схема ИАД позволяет на выборке примеров (описаний прецедентов наличия у пациентов диагностируемой патологии) и контрпримеров (описаний прецедентов отсутствия соответствующей патологии) строить такие логические функции (комбинации определенных значений определенных параметров из использованного в описаниях наблюдаемых прецедентов перечня), которые принимают значение *истина* на и только на примерах, а значение *ложь* – на и только на контрпримерах имеющейся обучающей выборки.

¹ Т.е. объяснениях, которые учитывают то, как устроен *реальный «мир»*, а не используемая для его анализа математическая модель.

Такие функции формируются в процессе анализа сходств описаний прецедентов из анализируемой БФ и представлены комбинациями *необходимых факторов* (конкретных значений конкретных параметров) возникновения диагностируемой патологии. *Непротиворечивость* формируемых таким путем ЭЗ обеспечена тем, что из всех порождаемых на примерах сходств (комбинаций факторов влияния) отбираются лишь те, которые не встречаются ни в одном из контрпримеров. (Действительно, ведь если конкретная комбинация факторов *влияния* «отвечает» за *наличие* целевого эффекта, то ее не должно быть ни в одном из контрпримеров – прецедентов из БФ, на которых целевой эффект не наблюдается). *Доказательность* персонализированного диагноза в этой схеме обусловлена его *неоспариваемостью* на имеющихся текущих эмпирических данных.

Характерные особенности предлагаемого подхода – существование и эффективная (полиномиально-быстрая в смысле сложности вычислений) проверяемость критерия существования функций предлагаемого вида для заданной выборки описаний прецедентов (условия *достоинства* текущей выборки для ее корректного разделения на примеры и контрпримеры функциями предлагаемого вида), экспоненциально быстро растущие с линейным увеличением размеров анализируемой БФ оценки емкости множеств таких функций, эффективная разрешимость условия сохранения непустоты множеств таких функций (*наследуемости* корректной отделимости примеров и контрпримеров а также выявления *устойчивых* эмпирических закономерностей - ЭЗ) при пополнениях исходной обучающей выборки описаниями новых прецедентов.

Практическая значимость развиваемого подхода демонстрируется² на примере задач диагностики так называемой псевдопрогрессии (*ПсП*) опухолей головного мозга человека. С *псевдопрогрессией* обычно ассоциируют состояние, характеризующееся фиксируемыми посредством МРТ/КТ *временными изменениями* в опухоли и перифокальных тканях, сопровождающееся увеличением объема опухоли и степени ее контрастирования, возникающее, как правило, в ранние сроки после лучевого лечения, *спонтанно регрессирующие* без дополнительного противоопухолевого лечения (см., например, Рис.1).

² См. https://itmcongress.ru/itm2019/agenda/section/21123_Seksionnoe_zasedanie_3_Sistemy_podderzhki_prinyatiyaresheniy_v_zdravookhraneni_ch2_Nauchnyy/ (слайды доклада на сайте Конгресса ИТМ-2019)

Характеризующие *псевдопрогрессию* изменения, фиксируемые объективными средствами, - это, прежде всего:

- усиление накопления контрастного вещества с расширением границы или без (T1 к/у),

- усиление отека в опухоли и перифокальных тканях (T2/Flair), а также
- увеличение объема опухоли за счет кисты и/или солидного компонента.

Псевдопрогрессия (на примере):

Пациент А., 6 лет, ПА ХСО, СРТ 54 Гр.

ПсП преимущественно солидного компонента



до лечения

через 5 мес

через 1 год

ПСЕВДОПРОГРЕССИЯ

Рисунок 1 – Пример динамики изменений опухоли у пациента в течении периода в 12 месяцев.

Исходным материалом для проведенных исследований являются накопленные примерно за 15 лет описания более 400 однотипных по гистологическому диагнозу опухоли пациентов НМИЦ нейрохирургии имени академика Н.Н.Бурденко МЗ РФ, характеризующихся соответствующими значениями почти двух сотен проблемно-ориентированных параметров. Демонстрируются примеры описывающих исследуемый эффект *ПсП* характеристических функций (в том числе – уточняющих представления об эффекте *ПсП*, формируемые традиционными статистическими средствами анализа данных), их содержательная интерпретация (охватывающая как факторы, которые обязаны присутствовать в описании прецедента наличия *ПсП*, так и факторы, которые должны отсутствовать при наличии *ПсП*), а также механизм фальсификации заключений, позволяющий отсекал артефакты при экстраполяции найденных ЭЗ на описания вновь анализируемых прецедентов (формировании персонализированных заключения о диагнозе у новых пациентов).

Проведенная таким образом апробация при решении задачи диагностики псевдопрогрессии (в том числе – дифференциальной диагностики *ПсП* с прогрессией опухоли, а также – выделения прогностических факторов развития *ПсП*) демонстрирует практическую актуальность предложенного нового подхода к анализу эмпирических данных, расширяющего доступные на сегодняшний день возможности доказательной медицины в части решения задач диагностического характера. Наиболее значимые характерные особенности развиваемого подхода – это:

- возможности надежно (доказуемо³ корректно) оперировать в том числе и *малыми (статистически незначимыми)* выборками описаний прецедентов;
- *персонализация* формируемых в процессе ИАД диагностических заключений и рекомендаций (аргументируемая точными детерминистскими средствами);

³ В понимании доказуемости как неоспариваемости на имеющихся эмпирических данных.

- возможности обеспечить единую (общую) схему ИАД для расширяющихся (постоянно пополняемых) новыми элементами выборок эмпирических данных;
- пошаговая интерпретируемость, неформальная (использующая содержательные термины и понятия исследуемой предметной области) объясняемость а также содержательная

фальсифицируемость диагностических заключений, формируемых в рамках предложенной схемы ИАД.

Предложенный подход к организации интеллектуального анализа эмпирических данных позволяет порождать практически значимые результаты при решении социально-значимых задач медицинской диагностики.

СПИСОК ЛИТЕРАТУРЫ:

1. Трунин Ю.Ю., Голанов А.В., Костюченко В.В., Галкин М.В., Хухлаева Е.А., Коновалов А.Н. Псевдопрогрессия доброкачественной глиомы на примере пилоидной астроцитомы среднего мозга. Клиническое наблюдение // Онкологический журнал: лучевая диагностика, лучевая терапия. – 2018. – Т. 1, № 1, с. 94-97
2. Трунин Ю.Ю., Голанов А.В., Костюченко В.В., Галкин М.В., Хухлаева Е.А., Коновалов А.Н. Увеличение объема пилоидной астроцитомы среднего мозга: рецидив или псевдопрогрессия? Клинич. наблюдение // Опухоли головы и шеи. – 2016. – Т.6, № 1, с. 68-75
3. Trunin Y., Golanov AV, Kostjuchenko VV, Galkin M.V., Konovalov A.N. Pilocytic Astrocytoma Enlargement Following Irradiation: Relapse or Pseudoprogression? // Cureus. – 2017. - V 9, № 2. – P.e1045-e1045
4. Финн В.К. Индуктивные методы Д.С.Милля в системах искусственного интеллекта // Искусственный интеллект и принятие решений. 2010. Ч.I: №3, С.3-21, Ч.II: №4, С. 14-40.
5. Забейайло М.И. О некоторых оценках сложности вычислений в ДСМ-рассуждениях // Искусств. интеллект и принятие решений. - 2015. – Ч.I: №1, С.3-17. – Ч.II: №2. С.3-17.
6. Забейайло М.И. О некоторых возможностях управления перебором в ДСМ-методе // Искусственный интеллект и принятие решений. – 2014. – № 3. – С.3-21.

UDC 616-71

9

Michael I. Zabezhailo

Federal research center "Computer science and control" (Russian Academy of Sciences), Moscow, Russia, m.zabezhailo@yandex.ru

Yrii Y. Trunin

"N.N.Burdenko National Medical Research Center of Neurosurgery" (Ministry of Health of the Russian Federation), Moscow, Russia, ytrunin@nsi.ru

On the evidence of medical diagnosis: intelligent data analysis of limited size samples of describing patients empirical data

Abstract. Some abilities to extend argumentation means of evidence-based medicine by mathematical tools of provability checking in open logical theories are discussed. Intelligent data analysis (IDA) scheme based on this theoretical planform is presented. The IDA-scheme is designed to solve diagnostic problems in medicine. Effectiveness and practical significance of the proposed approach is demonstrated by examples of diagnostics of human brain tumor pseudoprogression.

Keywords: Intelligent data analysis, evidence-based medicine, decision support systems in medicine, pseudoprogression of human brain tumor