

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
к.т.н. М.Н. Королевой на диссертационную работу
Чебанова Дмитрия Константиновича
«Прогнозирование ремиссии или рецидива у онкологических больных,
реализуемое в интеллектуальной системе»,
представленную на соискание ученой степени кандидата технических наук
по специальности 05.13.17 – «Теоретические основы информатики»

Актуальность темы диссертации

Тема представленной диссертационной работы бесспорно является актуальной в связи с необходимостью решения прикладных и фундаментальных задач в онкологии при помощи интеллектуального анализа данных. В клинической практике важной задачей является прогнозирование заболевания с целью планирования лечебного процесса и применения превентивной терапии. Открытие ключевых механизмов возникновения и развития злокачественной опухоли необходимо для создания новых препаратов и методов лечения. Решение подобных задач требует разработки соответствующих моделей, методов и алгоритмов, а также реализующих их программных систем. Такие системы могут позволить пользователю-исследователю анализировать регулярно накапливаемые экспериментальные данные и порождать новые знания о причинных механизмах возникновения изучаемых эффектов и явлений, действуя в режиме интеллектуального программного ассистента. Аналогичные задачи характерны не только для области онкологии и медицины, но и для других предметных областей, в которых присутствуют слабо формализованные данные, например, в технической диагностике и обеспечении информационной безопасности.

Новизна исследований и результатов

В рамках диссертационной работы Д.К. Чебанова получены следующие новые результаты:

1. Предложен способ представления разнородных данных, используемых в онкологии, для последующего интеллектуального анализа этих данных и получению новых знаний.

2. Разработан и реализован метод направленной навигации по диаграмме сходств описаний прецедентов с применением цепочек частичного порядка, который позволил при помощи последовательных шагов вычислять отдельные гипотезы и быстро оценивать их интерпретируемость. За счет такой управляемой навигации приближенный ДСМ-метод позволяет сфокусироваться, в первую очередь, на порождении содержательно интерпретируемых гипотез.

3. Получены новые знания в клинической и молекулярно-генетической онкологии, имеющие прикладную и фундаментальную значимость, которые, согласно анализу научных источников, до настоящего времени не удавалось получить иными методами анализа данных.

На основе указанных результатов автором разработана интеллектуальная система, позволяющая пользователю-исследователю в области онкологии проводить интеллектуальный анализ постоянно накапливаемых данных с целью прогнозирования заболевания и быстро находить в общем массиве гипотезы, обладающие интересующими эксперта свойствами.

Значимость и достоверность результатов работы

Практическая значимость диссертационной работы заключается в разработке интеллектуальной системы на базе ДСМ-метода автоматизированной поддержки научных исследований (АПИ), которая позволяет пользователю-исследователю предметной области проводить интеллектуальный анализ онкологических данных с целью прогнозирования заболевания и выявления ключевых биологических механизмов возникновения и развития злокачественной опухоли. Система позволяет оперировать исследовательскими данными, в том числе с учетом регулярного пополнения текущих эмпирических данных вновь получаемыми сведениями.

Реализовано прогнозирование онкологического заболевания «меланома кожи» – выявление пациентов, у которых ожидается наступление состояния «ремиссия» или «рецидив». Данная возможность позволяет клиническим специалистам НМИЦ онкологии им. Н.Н. Блохина выявлять пациентов с повышенным риском развития рецидива, которые классифицируются в группу усиленного мониторинга. Реализованная автором функция направленной навигации по диаграмме сходств, в рамках приближенного ДСМ-метода АПИ, позволяет быстро находить гипотезы, обладающие интересующими эксперта свойствами, что также позволяет проверять исходную базу фактов (БФ) на предмет наличия интерпретируемых гипотез. Функция протестирована экспертами НМИЦ онкологии им. Н.Н. Блохина и используется ими для решения исследовательских задач.

Достоверность результатов диссертационной работы Чебанова Д.К. подтверждается проведенными вычислительными экспериментами, использованием результатов в отчетах НИР, поддержанных РФФИ, а также публикациями результатов исследований на профильных конференциях и семинарах, в том числе международных. Основные результаты, представленные в диссертации, изложены в четырех статьях, опубликованных в рецензируемом научном журнале из списка ВАК, и одной публикации в тезисах международной конференции.

Содержание

Основной текст диссертации состоит из введения, пяти глав и заключения. Во введении обоснована актуальность темы диссертационного исследования, определены предмет исследований, цель и задачи работы, сформулированы основные результаты, охарактеризованы научная новизна и практическая значимость представляемой работы, описана апробация результатов, дана краткая информация о публикациях и структуре текста.

Первая глава «Обзор методов искусственного интеллекта для онкологических данных и постановка задачи» содержит описание актуальных задач в сфере исследования онкологических заболеваний, используемых типов представления знаний, логико-математический и программный инструментарий, постановку задачи диссертационной работы, а также обоснование выбора метода интеллектуального анализа данных.

Вторая глава «Описание ДСМ-метода АПИ» посвящена общему описанию ДСМ-метода автоматизированной поддержки исследований, реализации ДСМ-рассуждений, и постановке проблем реализации.

В третьей главе «Алгоритмическая реализация ДСМ-метода АПИ» представлен выбор способа программной реализации, параллельное выполнение процедур, оптимизацию работы с памятью, и результаты примененной оптимизации.

Четвертая глава «Описание интеллектуальной системы» включает описание компонентов разработанной ИС и ее графический интерфейс, который предназначен для взаимодействия пользователя с исходной БФ, настройками ДСМ-решателя и результатами исследования.

Пятая глава «Эксперименты по интеллектуальному анализу онкологических данных» включает описание БФ, представления клинических, генетических и иммунных данных и получаемых после ИАД знаний, постановку задачи проведения экспериментов и их результаты, а также обзор перспектив развития интеллектуальной системы, включая ее практическую значимость и универсализацию.

Замечания по работе

1. В оформлении текста работы присутствуют погрешности:
 - листинги никак не выделены, сделаны общим для текста работы шрифтом, что усложняет понимание;
 - подписи к таблицам и рисункам представлены не в соответствии с ГОСТ 2.105.
2. Список сокращений и условных обозначений не является обязательным элементом структуры диссертации, но его наличие могло бы значительно облегчить знакомство с текстом.
3. Присутствует ряд описок:
 - в параграфе 2.2. на стр. 40 в определении структур решающих предикатов для метода простого положительного сходства (для $a+$) и в упрощенном методе сходства-различия (для d_0) не хватает открывающей скобок. В первом случае общей закрывающей, во втором – общей открывающей. Эта описка повторяется в приложении В;
 - в тексте работы представлено два листинга 1 на стр. 57-59 и стр. 68;
 - параграфе 5.2.1 показатель «возраст» на стр. 74 – 58 лет, на стр. 75 – 57 лет;

– на стр. 73 приведена ссылка на приложение 1. Вероятно, имелось ввиду приложение А. Аналогично на стр. 88 приложение 2 и Б, 3 и В, соответственно.

4. Представленные в таблице 2 «Фрагмент исходных данных по иммунному статусу пациентов» данные требуют дополнительных пояснений. Неясно, в каких интервалах задаются показатели исходных данных, есть ли у них единицы измерений.

5. В параграфе 4.1 фигурирует словосочетание «устойчивость системы», при этом отсутствует пояснение, что именно подразумевается.

6. В параграфе 5.5 «Результаты экспериментов» на стр. 92 представлена гипотеза в некорректной математической записи: « $CD8 \in 39.9-54.1$ & $IRI \in 0.29-0.7$ », хотя в целом можно догадаться, что именно подразумевается автором.

7. В качестве научной новизны автор выделяет несколько положений, которые являются демонстрацией инженерного искусства, показывают квалификацию соискателя как разработчика ИТ-систем, но не несут в себе научную новизну. Например, «*впервые интеллектуальная система (ИС) на основе ДСМ-метода АПИ адаптирована для работы на промышленном вычислительном кластере, что дало существенное сокращение вычислительного времени, и продемонстрировало возможность запуска ДСМ-метода АПИ на высокопроизводительных суперкомпьютерах подобного класса*» – если ДСМ-метод АПИ работает на персональном компьютере, очевидно, его можно реализовать и на суперкомпьютере. Также «*в программной реализации ИС использован язык `python`, являющийся одним из наиболее распространенных, что потенциально расширяет круг исследователей, способных проводить адаптацию ИС под собственные нужды, а также позволяет проводить несложную интеграцию с другими модулями*» – сложно спорить с верным утверждением, но научной новизны в данном случае нет.

8. Для реализации ДСМ-метода при разработке ИС автором был использован прием обработки массивов *list comprehension*. Однако для языка *python*, на котором реализован метод, существуют библиотеки с открытым исходным кодом *NumPy* и *SkyPy*. Их использование увеличивает эффективность вычислений с матрицами.

9. Одним из результатов диссертационной работы заявлена разработанная ИС. Значит, разработанная программа является неотъемлемой частью работы, и её программный код должен быть представлен либо в самой работе в качестве приложения, либо ссылки на репозиторий. В тексте указано, что заявка на регистрацию программного продукта подана в Роспатент, но нет информации о планируемой лицензии, будет ли выложен программный код в публичный доступ или это коммерческий продукт.

Заключительная оценка

Диссертационная работа Чебанова Д.К. «Прогнозирование ремиссии или рецидива у онкологических больных, реализуемое в интеллектуальной

системе» является завершенной научно-квалификационной работой, носящей больше практический характер реализации разработанных ранее фундаментальных теоретических изысканий. Автореферат соответствует содержанию диссертации. Диссертационная работа выполнена автором самостоятельно, соответствует критериям, установленным Положением о присуждении ученых степеней, и пунктам 4 и 5 паспорта специальности 05.13.17 – «Теоретические основы информатики».

Автор работы – Чебанов Дмитрий Константинович – при должном уровне защиты положений по научной новизне может быть достоин присуждения ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Официальный оппонент

кандидат технических наук (05.13.17),
доцент кафедры РК-9 "Компьютерные системы
автоматизации производства"
МГТУ им. Н.Э. Баумана
Моб.: +79266295614
e-mail: maria.nik.koroleva@gmail.com



Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»
Адрес: 105005 Москва, 2-я Бауманская ул., д. 5, стр. 1
Телефон: 8 (499) 263-6391
Адрес электронной почты: bauman@bmstu.ru