

ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертационную работу Трофимова Юрия Владиславовича
«Методы и гетерогенная гибридная платформа многоуровневого ансамблирования
объяснений при комплексировании моделей машинного обучения», представленную
на соискание учёной степени кандидата технических наук по специальности:

1.2.1. – «Искусственный интеллект и машинное обучение»

Актуальность диссертационного исследования Трофимова Ю.В. определяется широким применением моделей искусственного интеллекта в задачах с высокой ценой ошибки, прежде всего в здравоохранении, где наряду с точностью решения требуется его обоснованность, проверяемость и воспроизводимость. С усложнением интеллектуальных систем и переходом к комплексированию нескольких моделей машинного обучения одного уровня интерпретации становится недостаточно: объяснения, получаемые различными методами, часто расходятся между собой, а их совместное использование не опирается на единый формальный аппарат. Разработка методов многоуровневого формирования, согласования и объединения объяснений с учётом неопределённости является необходимым условием построения доверенных систем искусственного интеллекта и систем поддержки принятия врачебных решений.

Научная значимость исследования Трофимова Ю.В. заключается в разработке целостного научно-методического и программно-алгоритмического аппарата объяснимого искусственного интеллекта, обеспечивающего переход от использования отдельных методов объяснения к их системной интеграции в едином вычислительном контуре. Созданная на основе разработанных методов гетерогенная гибридная платформа обеспечивает оркестрацию вычислений, автоматизированный выбор моделей и методов объяснения, формирование и сопоставимую оценку качества объяснений и апробирована на биомедицинских задачах анализа маммографических изображений, диагностики стенозов коронарных артерий и заболеваний органов грудной клетки. Практическая значимость подтверждается внедрением результатов работы в специализированное программное обеспечение системы поддержки принятия врачебных решений в составе программного продукта резидента фонда «Сколково» ООО «МАКАО ИТ», а также тремя свидетельствами о государственной регистрации программ для ЭВМ.

Научная новизна работы Трофимова Ю.В. заключается в следующем:

1. Предложен метод многоуровневого формирования объяснений в иерархически организованных системах ИИ, отличающийся представлением объяснимого контура в виде ориентированного графа узлов и переходов с формированием объяснительных объектов на его узлах и рёбрах и выделением признакового, модельного, интерпретируемого и результирующего уровней.
2. Предложен метод согласования уровней объяснения на основе операторов межуровневого переноса, мер расхождения и показателей согласованности, устанавливающий формально проверяемую связь между уровнями объяснения.
3. Разработан формальный аппарат агрегирования, объединения и комплексирования гетерогенных объяснений, отличающийся разделением способов обработки однородных и разнородных объяснительных объектов и сохранением их типа и происхождения в итоговом объяснении.
4. Предложен метод квантификации неопределённости объединённого объяснения, учитывающий расхождения между методами объяснения, повторными запусками и моделями, а для мультимодальных систем – согласованность объяснений различных модальностей.
5. Предложен протокол воспроизводимого выбора метода объяснения и метрический контур его оценки, отличающиеся совместным учётом качества, устойчивости, неопределённости, вычислительной стоимости и контекста применения.

Разработанные Трофимовым Ю.В. методы использованы при выполнении двух проектов Российского научного фонда (№ 22-71-10112 и проект № 22-17-00045), и проекта Государственного задания Министерства науки и высшего образования Российской Федерации (№ 124112200072-2 «Применение объяснительного искусственного интеллекта для интерпретации моделей машинного обучения»). Результаты работы также используются в ГБУЗ МО «Дубненская больница», Лаборатории информационных технологий им. М.Г. Мещерякова ОИЯИ и в учебном процессе Университета «Дубна».

Всего по результатам диссертационного исследования Трофимовым Ю.В. опубликовано 40 работ. В журналах из перечня ВАК РФ по специальности 1.2.1 – «Искусственный интеллект и машинное обучение» – 3, в журналах из перечня ВАК по иным специальностям – 12, в изданиях, индексируемых в международных базах Scopus и Web of Science – 10. Получены 3 свидетельства о государственной регистрации программ для ЭВМ.

В ходе работы над диссертацией Трофимов Ю.В. продемонстрировал умение собирать, анализировать и систематизировать актуальные научные и практические результаты в области исследования, самостоятельно формулировать цели и задачи исследования, а также высокую квалификацию в проектировании и программной реализации сложных интеллектуальных систем.

Диссертационная работа Трофимова Ю.В. «Методы и гетерогенная гибридная платформа многоуровневого ансамблирования объяснений при комплексировании моделей машинного обучения» является законченной самостоятельной научно-квалификационной работой, содержащей новые результаты. Основные результаты работы были доложены на 26 международных и всероссийских научных конференциях.

Работа удовлетворяет всем требованиям, предъявляемым ВАК к диссертациям, представленным на соискание учёной степени кандидата технических наук по специальности 1.2.1 – «Искусственный интеллект и машинное обучение». Диссертант Трофимов Ю.В. является квалифицированным специалистом по данной специальности и заслуживает присвоения ему учёной степени кандидата технических наук.

Ведущий научный сотрудник, отдел № 14
Федерального государственного учреждения
«Федеральный исследовательский центр
«Информатика и управление»
Российской академии наук»
к.ф.-м.н., доцент
моб.: +7 (910) 422-71-82
эл. почта: averkin2003@inbox.ru



Аверкин А.Н.

