

ОТЗЫВ

**на автореферат диссертации Хрисанфова Михаила Дмитриевича
«Поиск ошибок в хроматографических базах данных и моделирование свойств
идентифицируемых молекул методами машинного обучения»
на соискание ученой степени кандидата химических наук по специальности**

1.4.2. Аналитическая химия (химические науки)

Современный химический анализ все в большей степени опирается на обработку больших массивов данных, получаемых с помощью хромато-масс-спектрометрических систем. Эффективность такой обработки напрямую зависит от качества используемых референсных баз данных и от адекватности алгоритмов интерпретации экспериментальной информации. В этом контексте диссертационное исследование М.Д. Хрисанфова, посвященное разработке методов машинного обучения для поиска ошибок в хроматографических базах данных и прогнозирования свойств органических молекул, является безусловно актуальным. Работа находится на стыке современной аналитической химии, хемоинформатики и методов искусственного интеллекта, что соответствует мировым тенденциям развития науки.

Автореферат свидетельствует о том, что диссертационная работа представляет собой завершенное научное исследование, обладающее высокой научной новизной и практической значимостью. Крайне важным результатом является разработка оригинального подхода к поиску ошибок в хроматографических базах данных, основанного на механизме голосования нескольких независимых предсказательных моделей. Его применение позволило выявить тысячи потенциально ошибочных записей в широко применяемых базах данных, что дает возможность проводить аудит массивов хроматографических данных. Не менее интересным представляется решение, предложенное для предсказания масс-спектров электронной ионизации по структуре молекулы. Разработка таких инструментов позволяет существенно расширить возможности применения этого метода ионизации для детектирования и установления строения вновь синтезируемых соединений. Важно отметить, что разработанные автором программные решения для реализации предложенных

алгоритмов размещены в открытом доступе и могут применяться в любой лаборатории.

Автореферат изложен логично, грамотным научным языком и содержит необходимый иллюстративный материал. Основные результаты исследования опубликованы в 4 статьях в рецензируемых научных журналах, включая издания первого квартиля международных наукометрических баз данных и первого уровня Белого списка журналов, а также прошли апробацию на многочисленных российских и международных конференциях, что подтверждает их достоверность и признание в научном сообществе.

В то же время, при ознакомлении с авторефератом возникли некоторые вопросы и замечания, носящие, скорее, дискуссионный характер и не снижающие общей высокой оценки работы:

1. В автореферате не указан объём выборок, использованных для обучения и валидации разработанных нейросетевых моделей (EI2FP и NEIMS-PyTorch), а также не приведено распределение данных на тренировочную, валидационную и тестовую части. Для задач машинного обучения этот параметр является критическим, поскольку позволяет оценить устойчивость моделей и их способность к обобщению. Вероятно, эти сведения приведены в полном тексте диссертации, однако в автореферате их отражение было бы желательным.
2. Для демонстрации работы модели EI2FP:Lite автором приведены лишь два примера идентификации «неизвестных» соединений. Из текста автореферата неясно, насколько представительна тестовая выборка, каковы статистические показатели эффективности подхода в целом и каким образом выбирались соединения для иллюстрации.
3. При описании соединений в таблице 3 автореферата используются названия, не соответствующие номенклатуре.

Следует отметить, что высказанные замечания не затрагивают сути проведенного исследования и не противоречат сделанным в работе выводам и выносимым на защиту положениям. Таким образом, диссертационная работа Хрисанфова Михаила Дмитриевича «Поиск ошибок в хроматографических базах данных и моделирование свойств идентифицируемых молекул методами машинного обучения» по актуальности, научной новизне, практической значимости и объему материала полностью соответствует критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В.Ломоносова. Автор же диссертационного исследования несомненно заслуживает присуждения ученой степени кандидата химических наук по научной специальности 1.4.2 Аналитическая химия (химические науки).

Кандидат химических наук,
ведущий научный сотрудник лаборатории спектральных и хроматографических исследований Федерального государственного бюджетного учреждения науки Ордена Трудового Красного Знамени Института нефтехимического синтеза им. А.В.Топчиева Российской академии наук (ИНХС РАН)
Борисов Роман Сергеевич

28 августа 2026 года

Контактные данные:

Подпись сотрудника ИНХС РАН Р.С. Борисова удостоверяю:

Ученый секретарь ИНХС РАН

д.х.н., доцент

Костина Ю.В.

28 августа 2026 года