

### **Отзыв официального оппонента**

на диссертационную работу **Малыгина Ивана Вячеславовича**

**«Адаптация алгоритмов машинного обучения к геофизическим задачам с дефицитом данных»,**

представленную на соискание ученой степени кандидата технических наук по специальности 25.00.10 – Геофизика, геофизические методы поисков полезных ископаемых.

Геофизика – наука экспериментальная, и сопоставимость теории с реальностью возможна только с помощью достоверных, корректно проанализированных данных. Их обработка и интерпретация принадлежат к самым ответственным работам. Задачи с дефицитом данных особенно сложны. Условно их можно разделить на два категории. В первую входят задачи, анализирующие редкие события, для описания которых принципиально нет большого количества данных. Во вторую категорию входят задачи с пропуском данных, что обусловлено недостаточным количеством измерений, коротким промежутком времени наблюдений, большой протяженностью объектов, перерывами в регистрации данных и т.д. Этот, один из наиболее сложных, класс задач и выбрал Иван Вячеславович Малыгин для своей диссертационной работы. Для их решения он предлагает использовать алгоритмы машинного обучения. Машинное обучение – главное направление искусственного интеллекта, позволяющее настраивать параметры алгоритма обработки данных или перестраивать его архитектуру, основываясь на опыте (обучающей выборке). К сожалению применение методов машинного обучения не является стандартным подходом и они еще недостаточно широко используются в геофизике. Это дает основание утверждать, что научная проблема, сформулированная в диссертации, является важной и **актуальной**, влияющей в итоге на решение важнейших задач геофизики.

**Цель и задачи исследований** соответствуют существующим тенденциям развития вычислительной геофизики, от корректно проведенных теоретических

исследований до разработки компьютерных систем. Результаты достоверны, их ограничения на применение четко обозначены. Например, при построении трехмерных сейсмических образов изучаемого региона указано, что метод применим, если упругие параметры среды плавно меняются между станциями, т.е. между станциями отсутствуют субвертикальные разломы.

Автор применяет хорошо себя зарекомендовавшие, удобные в интерпретации методы машинного обучения, такие как метод k ближайших соседей, метод распознавания образов. Естественно, каждый применяемый метод требует адаптации к конкретной проблеме. Решаются задачи, относящиеся к разным областям геофизики. По данным инверсии приемных функций для северной части Балтийского щита построены карты поверхности Мохоровичича и слоя с низкими значениями скорости поперечных сейсмических волн  $V_S$ , составлена трехмерная сейсмическая модель коры и верхней мантии изучаемого региона. Разработан новый метод построения трёхмерного распределения проводимости среды по данным межскважинного электромагнитного просвечивания. Создана интеллектуальная система прогнозирования ледового заторообразования для рек Северная Двина и Лена. При этом получены количественные оценки влияния гидрометеорологических параметров на процесс ледового заторообразования, т.е. представлена возможность проверки гипотез влияния тех или иных факторов на состояния процессов. С использованием машинного обучения эти задачи решались впервые. На мой взгляд, **новизна** заключается не только в том, что известные методы применены к решению новых задач, но и в том, что диссертационная работа раскрывает возможности и особенности машинного обучения, в том, что показана возможность применения машинного обучения к решению практически любых геофизических задач с дефицитом данных.

**Обоснованность** научных положений, выводов и рекомендаций подтверждается успешным решением конкретных задач. В свою очередь, для каждой конкретной задачи достоверность результатов подтверждается сравнением с выводами работ других авторов. Так, например, после построения карты поверхности Мохоровичича, диссертант соотносит ее с ранее полученными

результатами по профилям. Полученные в рамках настоящего исследования методом kNN результаты хорошо соответствуют предыдущим результатам и позволяют выявить локальные неоднородности. Для задачи прогнозирования ледового заторообразования на реках валидация разработанной системы проведена на новых, недоступных в момент первоначальной разработки, данных.

**Значимость полученных результатов** для науки состоит в расширении применимости существовавших методов исследований, исследовании свойств и характеристик среды при решении конкретных задач, создании теоретических предпосылок проектирования систем для других геофизических приложений при дефиците данных. Практическая значимость представлена решением прикладных задач пространственной интерполяция (двумерной и трехмерной) и классификации на основе временных рядов. Помимо исследований, представленных в работе, разработанный метод интерполяции геофизических данных алгоритмами машинного обучения применим для широкого круга задач геофизических измерений. Интеллектуальная система прогнозирования заторообразования может быть применена в любых регионах, а также для прогнозирования других опасных природных явлений с аналогичной структурой исходных данных, в том числе для проверки степени влияния тех или иных факторов на трудноформализуемый процесс в условиях дефицита данных.

Диссертация построена по традиционному принципу и состоит из введения, 3-х глав собственных исследований, заключения, списка использованной литературы, содержащего 168 наименований, и 4-х приложений. Работа иллюстрирована 16-ю таблицами и 39-ю рисунками. Обзор литературы дается в тексте, для каждой задачи отдельно. Диссертация написана грамотным техническим языком, изложение логично и последовательно. Содержание диссертации соответствует поставленным целям и задачам исследования. Автор квалифицированно подошел к построению алгоритмов машинного обучения и успешно реализовал их в виде комплексов программ, о чем имеются Свидетельства о государственной регистрации программ для ЭВМ.

Диссертант корректно использует известные научные методы обоснования полученных результатов, выводов и рекомендаций. Автореферат соответствует основному содержанию диссертации, отражает ее структуру и положения, выносимые на защиту. Основные научные результаты достаточно полно опубликованы в рецензируемых научных изданиях.

Существенных недостатков работа не содержит, тем не менее, необходимо сделать некоторые **замечания** по её содержанию.

1. Поскольку геофизика – широчайший комплекс научных дисциплин, полно обрисовать ситуацию дефицита данных в геофизических задачах в подобном исследовании не представляется возможным. Также невозможно полно и подробно описать все алгоритмы, применяемые в настоящее время для их решения. По-видимому стоило ограничиться классификацией задач с дефицитом данных с точки зрения оптимальных алгоритмов их обработки.

2. Диссертация была бы более полной при сравнении различных алгоритмов и их модификаций.

3. В Первую главу, посвященную в основном теоретическим положениям, включено описание исследования, направленного на поиск кимберлитовых тел методом построения трёхмерного распределения проводимости среды по данным межскважинного электромагнитного просвечивания. Имело бы смысл выделить эту конкретную задачу в отдельную главу.

4. В Первой главе четко, по пунктам, представлен метод адаптации алгоритмов машинного обучения к геофизическим приложениям с дефицитом данных. Однако во Второй и Третьей главе автор не всегда придерживается этого плана. Например, в конкретных задачах выбор функционала качества недостаточно обосновывается.

5. При оформлении диссертации и автореферата не удалось избежать некоторой путаницы при обозначении величин и при использовании терминов. Так, например, известный в геофизике термин «скейлинг» приобретает у автора несколько иной смысл.

**Заключение.** Отмеченные недостатки не влияют на главные теоретические и практические результаты работы. Диссертация Малыгина Ивана Вячеславовича

«Адаптация алгоритмов машинного обучения к геофизическим задачам с дефицитом данных» выполнена на актуальную тему, обладает научной новизной и содержит решение задачи применения и адаптации алгоритмов машинного обучения к геофизическим задачам с дефицитом данных, имеющей существенное значение для российской науки и практики. Полученные автором результаты являются обоснованными и достоверными.

*Диссертация соответствует критериям, установленным п. 9 Положения о присуждении учёных степеней (утверждено постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842) для учёной степени кандидата наук, а её автор, И.В. Малыгин, достоин присуждения учёной степени кандидата технических наук по специальности 25.00.10 – Геофизика, геофизические методы поисков полезных ископаемых.*

**Официальный оппонент:** старший научный сотрудник

*Лаборатории прямых и обратных задач, изучения очага землетрясений*

*Федерального государственного бюджетного учреждения науки*

*Института теории прогноза землетрясений и математической геофизики*

*Российской академии наук (ИТПЗ РАН)*

*кандидат физико-математических наук Кислов Константин Викторович*

*117997, Москва, улица Профсоюзная, 84/32*

*[kvkislov@yandex.ru](mailto:kvkislov@yandex.ru), тел. +7(909)9571306*

**Я, Кислов Константин Викторович, даю согласие на включение своих персональных данных в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.**

  
К.В. Кислов

10 августа 2022 года

Подпись официального оппонента заверяю

Заместитель директора по общим вопросам



  
Каберов Ваис Харисович