

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.073.05  
НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО УЧРЕЖДЕНИЯ  
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР «ИНФОРМАТИКА И  
УПРАВЛЕНИЕ» РОССИЙСКОЙ АКАДЕМИИ НАУК» ПО ДИССЕРТАЦИИ  
НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА НАУК

аттестационное дело № \_\_\_\_\_

решение диссертационного совета от 24 декабря 2020 г. № 9

О присуждении Ирхину Илье Александровичу, гражданину Российской Федерации, ученой степени кандидата физико-математических наук.

**Диссертация** «Единственность матричного разложения и сходимость регуляризованных алгоритмов в вероятностном тематическом моделировании» по специальности 05.13.17 — «Теоретические основы информатики» принята к защите 22 октября 2020 г., протокол №5, диссертационным советом Д002.073.05 на базе Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», 119333, г. Москва, ул. Вавилова, д. 44/2, созданным на основании приказа Министерства Образования и Науки Российской Федерации №783/нк от 24.06.2016 г.

Заседание диссертационного совета проводилось в удаленном интерактивном режиме. Основание — приказ ФИЦ ИУ РАН № 3-78 от 11.12.2020 г.

**Соискатель** Ирхин Илья Александрович, 1993 года рождения, в 2016 году окончил Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (государственный университет)» по направлению подготовки 01.04.02 Прикладная математика и информатика.

В 2020 году закончил очную аспирантуру Федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)».

С 21.09.2020 был прикреплен к аспирантуре Федерального исследовательского центра «Информатика и управление» Российской Академии Наук для сдачи кандидатского экзамена.

Диссертация выполнена на кафедре «Анализ данных» физтех-школы прикладной математики и информатики Федерального государственного автономного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)».

Соискатель в настоящее время работает в ООО «Яндекс.Такси Технологии» руководителем отдела машинного обучения и анализа данных.

**Научный руководитель** — доктор физико-математических наук, профессор РАН Воронцов Константин Вячеславович. Работает в ФГАОУ ВО «Московский физико-технический институт (национальный исследовательский университет)» в должности руководителя лаборатории машинного интеллекта.

#### **Официальные оппоненты**

Елизаров Александр Михайлович, доктор физико-математических наук, профессор кафедры программной инженерии Высшей школы информационных технологий и интеллектуальных систем ФГАОУ ВО «Казанский (Приволжский) федеральный университет»,

Фильченков Андрей Александрович, кандидат физико-математических наук, доцент факультета «Информационных технологий и программирования» ФГАОУ ВО «Национальный исследовательский университет ИТМО»

**дали положительные отзывы на диссертацию.**

**Ведущая организация** Федеральное государственное бюджетное учреждение науки Институт системного программирования им. В.П. Иванникова Российской академии наук (ИСП РАН) **в своем положительном отзыве**, подписанном Турдаковым Денисом Юрьевичем, к.ф.-м.н., зав.отделом информационных систем указала, что диссертация является законченной научно-квалификационной работой, содержащей новые научные результаты в области семантического анализа текстов. Все результаты, выносимые на защиту, обоснованы и подтверждены в ходе вычислительных экспериментов. В заключении отмечается,

что диссертация соответствует всем критериям, установленным Положением о присуждении ученых степеней, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук, а ее автор — Ирхин Илья Александрович заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.17 — «Теоретические основы информатики».

Соискатель имеет 3 опубликованных работы по теме диссертации, 3 из которых опубликованы в рецензируемых научных изданиях, рекомендованных ВАК. **Наиболее значимые публикации:**

1. Селезнев Н.К., Ирхин И.А., Кантор В.В. Автоматическое извлечение атрибутов водителя из логов мобильного приложения такси // Труды МФТИ. — 2018. — Т. 10, № 3. — С. 5—15.
2. Ирхин И.А., Воронцов К.В. Сходимость алгоритма аддитивной регуляризации тематических моделей // Труды Института математики и механики УрО РАН. — 2020. — Т. 26, № 3. — С. 57—68.
3. Дербаносов Р. Ю., Ирхин И.А. Проблемы устойчивости и единственности стохастического матричного разложения // Журнал вычислительной математики и математической физики. — 2020. — Т. 60, № 3. — С. 19—28.

Недостовверных сведений в работах, опубликованных соискателем ученой степени, не выявлено. В коллективных публикациях автору принадлежат те их части, которые легли в основу положений диссертации, выносимых на защиту.

**Выбор официальных оппонентов и ведущей организации** обосновывается тем, что оппоненты и сотрудники ведущей организации являются известными специалистами в области разработки и исследования моделей и алгоритмов анализа данных и обработки естественного языка, имеют публикации в авторитетных научных журналах в соответствующих сферах исследований и могут определить научную и практическую значимость диссертации.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

**доказаны** теоремы о достаточных условиях сходимости EM-алгоритма для построения тематических моделей с аддитивной регуляризацией;

**предложена** модификация EM-алгоритма, основанная на доказанных теоремах, ускоряющая сходимость итерационного процесса;

**доказана** теорема о достаточных условиях единственности стохастических матричных разложений и выявлена важность разреженности модели для обеспечения единственности разложения в точке сходимости;

**предложен** метод разреживания тематической модели, практически не ухудшающий качество модели (в смысле правдоподобия или перплексии).

**Теоретическая значимость** исследований обусловлена тем, что:

полученный комплекс результатов является первым строгим теоретическим обоснованием численных методов построения аддитивно регуляризованных тематических моделей (ARTM);

полученные теоретические результаты позволили диссертанту не только обосновать уже используемые численные методы ARTM, но и предложить ряд нетривиальных модификаций, улучшающих сходимость EM-алгоритма, разреженность получаемых моделей и другие критерии качества моделей.

**применительно к проблематике диссертации результативно использованы** методы доказательства сходимости обобщённых EM-алгоритмов, методы доказательства единственности неотрицательных матричных разложений, методы теории вероятностей, линейной алгебры, численные методы оптимизации;

**Значение полученных соискателем результатов исследования для практики** подтверждается тем, что:

**разработаны и реализованы** все предложенные модификации EM-алгоритма для аддитивной регуляризации тематических моделей;

**создана** и размещена в открытом доступе библиотека функций на языке Python с программной реализацией всех предложенных модификаций EM-алгоритма для аддитивной регуляризации тематических моделей;

некоторые из предложенных в работе алгоритмов **реализованы** в библиотеке с открытым кодом TopicNet.

## Оценка достоверности результатов исследования выявила:

для экспериментальных работ использовались стандартные для данной области наборы данных, критерии качества и открытое программное обеспечение, в тексте диссертации представлены все необходимые детали для воспроизведения полученных результатов;

теоретические результаты обоснованы строгими математическими доказательствами теорем.

**Личный вклад соискателя.** Работа над теоремой о единственности стохастического матричного разложения выполнена совместно с Дербаносовым Р. Ю. Соискателем была доказана основная лемма, реализованы и проведены все эксперименты. Теоремы о сходимости алгоритма аддитивной регуляризации тематических моделей, модификации этого алгоритма, доказательства их теоретических свойств, реализация и проведение экспериментов выполнены соискателем лично.

На заседании 24 декабря 2020 года диссертационный совет принял решение присудить Ирхину Илье Александровичу ученую степень кандидата физико-математических наук.

При проведении открытого голосования диссертационный совет в количестве 21 человек, из них 7 докторов наук по профилю рассматриваемой диссертации, участвовавших в заседании, из 30 человек, входящих в состав совета, дополнительно введены на разовую защиту 0 человек, проголосовали за 21, против 0.

И.о. председателя диссертационного совета Д 002.073.05,

доктор технических наук



Матвеев И.А.

Ученый секретарь диссертационного совета Д 002.073.05,

кандидат технических наук

Рейер И.А.

« 24» декабря 2020г.