

Отзыв научного руководителя

о диссертационной работе Ирхина Ильи Александровича
«Единственность матричного разложения и сходимость регуляризованных алгоритмов
в вероятностном тематическом моделировании»,
представленной на соискание ученой степени кандидата физико-математических наук
по специальности 05.13.17 — «Теоретические основы информатики»

Вероятностное тематическое моделирование — одно из направлений обработки естественного языка, активно развивающееся с конца 90-х годов. Тематическая модель выявляет тематику коллекции текстовых документов, определяя, к каким темам относится каждый документ, и какие слова образуют каждую тему. Данная задача сводится к низкоранговому стохастическому матричному разложению и является в общем случае некорректно поставленной по Адамару. Диссертационное исследование И. А. Ирхина выполнено в рамках теории аддитивной регуляризации тематических моделей (ARTM), в которой множество дополнительных требований к модели формализуется с помощью взвешенной суммы оптимизационных критериев, называемых регуляризаторами. Теоретически это должно улучшать устойчивость и приводить к единственному решению, однако свойства единственности и устойчивости в рамках подхода ARTM до сих пор были слабо изучены.

Целью диссертационной работы И. А. Ирхина было доказательство теоретических свойств алгоритма ARTM, а также поиск возможных улучшений метода. Поставленные задачи выполнены диссертантом в полном объёме. Были доказаны теоремы о сходимости итерационного процесса ARTM и теорема о единственности стохастического матричного разложения. Также на основе полученных теоретических результатов были предложены и реализованы модификации постановки оптимизационной задачи ARTM и EM-подобного алгоритма оптимизации, которые показали заметное улучшение качества получаемых тематических моделей в экспериментах на текстовых коллекциях.

И. А. Ирхиным был проделан большой объём работы: сделан обзор современной научной литературы, проведено теоретическое исследование, полученные теоретические результаты оформлены в виде лемм и теорем, разработаны и реализованы модификации EM-алгоритма для ARTM и проведены эксперименты, подтверждающие их применимость на практике.

Диссертационная работа И. А. Ирхина «Единственность матричного разложения и сходимость регуляризованных алгоритмов в вероятностном тематическом моделировании» является законченной самостоятельной научно-исследовательской

работой, содержащей новые математические результаты. Основные результаты диссертации докладывались на российских и международных научных конференциях и научных семинарах. Следует также отметить педагогическую работу И. А. Ирхина – преподавание машинного обучения и математической логики и теории алгоритмов на факультете ФИВТ МФТИ.

Работа удовлетворяет всем требованиям, предъявляемым ВАК к диссертациям, представляемым на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 — «Теоретические основы информатики». Диссертант И. А. Ирхин является квалифицированным специалистом по данной специальности и заслуживает присвоения ему ученой степени кандидата физико-математических наук.

24 сентября 2020 г.

Руководитель лаборатории
машинного интеллекта МФТИ,
профессор РАН, д.ф.-м.н.
Моб. +7 916 333 71 69
eMail: k.v.vorontsov@phystech.edu

Воронцов К. В.

ЗАБЕРЯЮ
УЧЕНЫЙ СЕКРЕТАРЬ

МФТИ
Ю. И. СКАДОВНИКОВ



В. Калку