

## Отзыв научного руководителя

Стрижова Вадима Викторовича

о диссертационной работе Кузьмина Арсения Александровича  
«Иерархическая классификация коллекций документов»,  
представленной на соискание ученой степени кандидата физико-математических наук  
по специальности 05.13.17 – теоретические основы информатики

Целью диссертационной работы А.А. Кузьмина является разработка и теоретическое обоснование методов иерархической классификации коллекций документов, автоматически структурирующих документы в виде иерархий тем и оптимизирующих иерархии, выявляя в них тематические несоответствия. Рассматриваются частично размеченные коллекции документов с экспертно заданными древовидными структурами тем. Решением задачи классификации для каждого неразмеченного документа является ранжированный список кластеров нижнего уровня иерархии согласно их релевантности данному документу. Получены следующие основные результаты: предложена иерархическая взвешенная функция сходства кластеров; построен оператор релевантности, ранжирующий кластеры в порядке убывания сходства с заданным документом; предложена энтропийная модель оценки важности слов при классификации; предложена вероятностная модель иерархической классификации и способ оптимизации ее параметров и гиперпараметров; разработан комплекс программ для иерархической классификации тезисов крупной конференции и веб-сайтов компаний индустриального сектора.

А.А. Кузьминым проделан большой объем как теоретической, так и экспериментальной работы. Экспериментальная часть работы содержит сравнение предложенных методов с существующими аналогами и рекомендации по построению экспертных систем для классификации частично размеченных коллекций документов, в которых экспертную структуру тем можно представить в виде дерева.

А.А. Кузьмин закончил бакалавриат и магистратуру МФТИ с отличием. В аспирантуре он продолжил исследования алгоритмов текстовой классификации и получил значимые теоретические и экспериментальные результаты, продемонстрировав хорошее владение математическим аппаратом, самостоятельность, умение ставить новые цели и достигать их. А.А. Кузьмин вел активную научную деятельность: принимал участие в крупных международных конференциях по исследованию операций и анализу данных, публиковал полученные результаты в рецензируемых журналах, в том числе в журналах, рекомендованных ВАК, руководил студенческими исследованиями.

Диссертация А.А. Кузьмина является серьёзной законченной научно-квалификационной работой в области разработки и исследования алгоритмов анализа данных. Она удовлетворяет всем требованиям, предъявляемым к диссертациям на соискание учёной степени кандидата физико-математических наук по специальности 05.13.17 - «Теоретические основы информатики», а сам Кузьмин А.А. заслуживает присуждения ему учёной степени кандидата физико-математических наук по указанной специальности.

Научный руководитель  
доктор физико-математических наук,

Стрижов Вадим Викторович

Наименование организации – места работы: Вычислительный центр им. А.А. Дородницына  
Российской академии наук Федерального исследовательского центра «Информатика и  
управление» Российской академии наук

Структурное подразделение: Отдел интеллектуальных систем

Должность: научный сотрудник

Адрес электронной почты: strijov@ccas.ru

Телефон: 8 (499) 135-41-63

Адрес организации: 119333, г. Москва, ул. Вавилова, д. 40

25 января 2017 г.

Подпись Стрижова В.В. и сведения заверяю

