

ОТЗЫВ

об автореферате диссертации Бондарчука Дмитрия Вадимовича «Алгоритмы интеллектуального поиска на основе метода категориальных векторов», представленной на соискание учёной степени кандидата физико-математических наук по специальности 05.13.17 – теоретические основы информатики

Диссертантом Д.В. Бондарчуком предлагается новый метод интеллектуального анализа данных, который на любой запрос пользователя дает пользователю непустой ответ, отсортированный по степени релевантности запросу пользователя. Алгоритм обучения состоит из этапов обработки обучающей выборки текстов с использованием стеммера и определении частот вхождений термов в документы, уменьшения количества термов с использованием сингулярного разложения матрицы корреспонденций термов, вычисления категориальных векторов документов базы данных вакансий и пользовательского запроса, вычисления коэффициентов близости между пользовательским запросом и данными базы, сортировки по убыванию и выбора части первых элементов. В качестве модели представления выбрана векторная модель, в которой каждый текстовый документ из коллекции представляется, как вектор в векторном пространстве.

Разработаны модель образа текстового документа и соответствующий метод отображения текста в семантическое пространство, обеспечивающие компактное представление документа в оперативной памяти на основе матрицы корреспонденций термов, которая подвергается ортогональному разложению. Предложен метод перевзвешивания термов векторной модели с помощью вычисления их семантической взаимосвязи друг с другом на основе авторской адаптации алгоритма Леска.

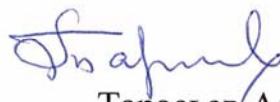
На основе разработанных методов и подходов реализован алгоритм подбора рекомендаций. Проведены вычислительные эксперименты, подтверждающие более высокую эффективность разработанного

алгоритма по сравнению с существующими, что имеет несомненную практическую ценность данного исследования.

Исходя из представленного автореферата и публикаций можно сделать вывод о достаточной обоснованности полученных результатов, вследствие использования математически строгих доказательств и подкрепления теоретических материалов вычислительными экспериментами.

Полагаю, что диссертационное исследование Д.В. Бондарчука является завершенным научным исследованием, имеющим существенное значение в области интеллектуального анализа текстовых данных. Работа отвечает требованиям Положения о порядке присуждения ученых степеней, предъявляемым к кандидатским диссертациям, а её автор Дмитрий Вадимович Бондарчук заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 05.13.17 – теоретические основы информатики.

Зав. отделом динамических систем
Института математики и механики
им. Н.Н. Красовского (ИММ)
Уральского отделения Российской
академии наук,
доктор физико-математических
наук



Тарасьев Александр Михайлович

e-mail: dir-info@imm.uran.ru,
tam@imm.uran.ru
Раб. телефон: +7 (343) 374-83-32,
+7 (343) 362-81-73
адрес организации:
620990, Российская
Федерация, г. Екатеринбург, ул. Софьи
Ковалевской, д. 16

Подпись А.М. Тарасьева заверяю:
Ученый секретарь ИММ УрО РАН,
к.ф.-м.н.



О.Н. Ульянов