

ОТЗЫВ ЗАРУБЕЖНОГО НАУЧНОГО КОНСУЛЬТАНТА

о диссертационной работе Якунина Кирилла Олеговича «Разработка моделей и методов сбора, анализа и классификации медиа-публикаций на базе методов обработки естественных языков», представленной на соискание степени «доктор философии» (PhD) по специальности «6D070400 - Вычислительная техника и программное обеспечение»

Классификация текстов является трудной задачей, требующей использования сложных математических и статистических моделей, а также больших объемов размеченных экспертами примеров. При этом векторное представление текстов является одной из ключевых проблем, поскольку практика показывает, что эффективные векторные представления позволяют добиться высокого качества классификации даже при использовании относительно простых моделей машинного обучения.

В своей работе Якунин К.О. предложил модели и методы, позволяющие обучать классификаторы на базе так называемых тематических векторизаций текстов при небольшом объеме ручной экспертной разметки. Разметку предлагается проводить не на уровне отдельных текстов, а на уровне либо скрытых латентных структур корпуса (топиков/тем), либо на уровне какой-либо релевантной мета-информации о публикациях (источник, автор, показатели активности и т.п.).

В ходе проведения диссертационного исследования Якунин К.О. проделал анализ известных подходов и методов решения поставленных задач, выявил основные ограничения современных подходов. Автор смог предложить подходы к решению поставленных задач, позволившие получить ряд прикладных и исследовательских результатов.

В диссертации получены следующие научные результаты:

1. Предложен метод векторизации текстовых документов с помощью тематической модели BigARTM.
2. Предложен метод оценки тематического межкорпусного дисбаланса для самообучения классификационной модели.
3. Предложена методика многофакторной оценки социальной значимости публикации.
4. Предложена методика многокритериальной оценки масс-медиа ММА на базе байесовской системы агрегации, метода анализа иерархий (АНР) и тематического моделирования.

Практическая ценность работы заключается в том, что на основе предложенных методов и алгоритмов была разработана информационная система, позволяющая проводить автоматический мультикритериальный анализ медиа-пространства. Система была внедрена в МОН РК в 2020г.

Все результаты автором получены лично. Научные руководители принимали участие в формулировке задач, согласовании плана проведения исследования, а также, в оценке результатов. Диссертация написана доходчиво, грамотно и аккуратно оформлена. Аннотация соответствует основному содержанию диссертации.

В ходе диссертационного исследования автором было опубликовано 15 научных трудов, из которых 5 статей опубликованы в журналах, входящих в базы Scopus и Thomson Reuters (2 статьи – Q1, 1 статья – Q2, 2 статьи – Q3), одна из них опубликована в изданиях, рекомендованных Комитетом по контролю в сфере образования и науки МОН РК, 3 статьи в журналах имеющих CiteScore (2 статьи - CiteScore 17%, 1 статья – CiteScore 69%) в Scopus без присвоенного квартиля, 2 статьи в журналах, входящих в РИНЦ (импакт-фактор 0.482 и 1.385), 6 статей опубликованы в сборниках международных научно-практических конференций. Оформлено 4 авторских свидетельства на результаты работы.

Якунин Кирилл прошёл научную стажировку в 2021 году на базе Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр информационных и вычислительных технологий» (ФИЦ ИВТ) (Новосибирск, РФ) под моим руководством.

На основании вышеизложенного считаю диссертационную работу Якунина Кирилла Олеговича на тему «Разработка моделей и методов сбора, анализа и классификации медиа-публикаций на базе методов обработки естественных языков» завершённой, удовлетворяющей требованиям, предъявляемым к диссертациям на соискание степени доктора философии (PhD) по специальности 6D070400 - Вычислительная техника и программное обеспечение, и рекомендую к защите.

Зарубежный научный консультант
д.т.н., доцент



В.Б. Барахнин
17.05.2021г.

Подпись В.Б.Барахнина заверяю
Ученый секретарь ФИЦ ИВТ
к.т.н.



А.С. Рылов