

ОТЗЫВ

научного консультанта на диссертационную работу **Мухсиной Куралай Женисбековны** «Разработка системы анализа многоязычной текстовой информации на основе машинного обучения», представленную на соискание ученой степени доктора философии (PhD) по специальности «6D070400 – Вычислительная техника и программное обеспечение».

Диссертационная работа Мухсиной К.Ж. посвящена проблеме развития моделей, методов и алгоритмов обработки текстовой информации на основе машинного обучения и применения этих методов для автоматизации анализа многоязычной текстовой информации.

Анализ научных исследований в области компьютерной лингвистики, интеллектуального анализа и искусственного интеллекта, связанных с включением казахского языка в многоязычные проекты, показывает, что существующих в данном научном направлении решений не достаточно для удовлетворения на качественном уровне существующих потребностей по разработке систем автоматической обработки текстов казахского, русского и английского языков.

Многоязычные приложения NLP используют только грамматический этап обработки текста, а семантический анализ текста или анализ смысла естественного языка, по-прежнему остается одной из основных проблем теории искусственного интеллекта и компьютерной лингвистики. Для использования методов машинного обучения при семантической и грамматической обработке многоязычной информации необходимы заранее созданные грамматически и семантически размеченные корпуса каждого естественного языка.

В связи с этим, все вышеперечисленное и обуславливает актуальность диссертационной работы, посвященной комплексному исследованию и решению проблемы анализа многоязычной текстовой информации на основе машинного обучения.

Таким образом, поставленные в диссертационной работе задачи исследования полностью решены. Программная реализация и проведенная оценка качества работы системы, извлекающей знания в виде триплета фактов из текста, и системы, определяющей семантическую близость текста к узкоспециализированной тематике, базирующихся на разработанных в диссертационном исследовании моделях и алгоритмах, показывает хорошие результаты.

Постановка проблемы диссертационного исследования носит прикладной характер. Диссертационная работа подготовлена в рамках реализации научного проекта МОН РК «Методы и модели поиска и анализа криминально значимой информации в неструктурированных и слабоструктурированных текстовых массивах»– AP05131073 (2018-2020 гг.).

Диссертационное исследование Мухсиной К.Ж. представляет научный и практический интерес, вносит вклад в дальнейшее становление, как теории искусственного интеллекта, так и компьютерной лингвистики.

Диссертационная работа представляет собой научный труд, имеющий внутреннее единство. Содержание работы соответствует логике изложения и полностью раскрывает тему исследования, структурные элементы диссертации не вызывают возражений. Работа оформлена в соответствии с требованиями по подготовке диссертаций. Результаты, полученные в ходе исследования, достоверны и обоснованы. Достоверность и обоснованность результатов исследования определяется практической реализацией программного продукта, представляющего динамически пополняемый узкоспециализированный выровненный параллельный многоязычный корпус.

В целом, выполненная диссертация может быть охарактеризована как самостоятельное и завершенное исследование, имеющее научную и практическую

ценность и отвечающее требованиям, предъявляемым к докторским работам по данной специальности. Актуальность, новизна и обоснованность выводов, теоретическая и практическая значимость диссертационного исследования позволяют заключить, что диссертационная работа Мухсиной К.Ж. на тему «Разработка системы анализ многоязычной текстовой информации на основе машинного обучения» может быть рекомендована к публичной защите на соискание степени доктора философии (PhD) по специальности «6D070400 – Вычислительная техника и программное обеспечение».

Научный консультант,
PhD, ассоциированный профессор Мамырбаев О.Ж.

