

РЕЦЕНЗИЯ

на выпускную квалификационную работу обучающегося СПбГУ

Новосёловой Анастасии Максимовны

по теме Исследование методов автоматического реферирования текстов

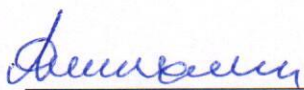
Выпускная квалификационная работа Новосёловой Анастасии Максимовны посвящена решению задачи автоматического реферирования текстов. Данная задача является актуальной на сегодняшний день, поскольку при работе с большими по объему документами порой необходимо иметь возможность ознакомления с краткой аннотацией для того, чтобы иметь возможность оперативно принять решение о дальнейшей работе с ним. Такая возможность востребована прежде все в системах электронного документооборота, которые отвечают за обеспечение работы над юридическими документами в коммерческих компаниях и государственных организациях. Кроме этого, алгоритмы для автоматического построения рефератов могут быть использованы для сокращения текстов новостных сообщений с целью сокращения времени ознакомления с ними.

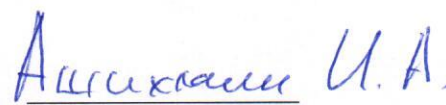
Рецензируемая работа выполнялась с использованием набора новостных данных, предоставленных компанией Digital Design. В первых двух главах работы приводится постановка задачи и описание используемых алгоритмов. Автором рассмотрены следующие подходы к решению проблемы построения рефератов: TextRank, k-means, LSA. Помимо этого, во второй главе приводятся различные метрики, позволяющие оценить качество работы разных подходов. Третья глава посвящена представлению результатов работы различных алгоритмов и приводится оценка их качества на основании описанных выше метрик. В результате исследований автор приходит к выводу, что наилучший результат дает метод TextRank и делает предложение по улучшению его работы с помощью модификации способа вычисления степени схожести предложений. Реализация алгоритма TextRank с использованием меры схожести в виде расчета длины наибольшей общей подпоследовательности позволила немного улучшить результат, полученный изначально.

В целом Анастасия Максимовна с поставленной задачей справилась. По результатам работы подготовлен доклад на конференцию по проблемам информатики СПИСОК-2017. Считаю, что представленная работа заслуживает оценки "отлично".

Руководитель научной лаборатории
Компания Digital Design
Ашихмин Илья Алексеевич

«30» мая 2017 г.


Подпись


ФИО