



Отчет об оценке

докторской диссертации под названием

«Разработка метода для информативного экстрактивного реферирования научных текстов на английском языке»

представленный Ахметовым Искандером Рафаиловичем

1. Подробный отчет и комментарии

а) Цель и сфера применения

Диссертация посвящена очень важной области вычислительной техники: обработке естественного языка. Эта дисциплина создает методологии, позволяющие компьютерам осмысленно обрабатывать человеческий язык. Среди наиболее важных приложений этой технологии - поиск информации и автоматическое реферирование текстов.

Задача автоматического реферирования научных текстов очень важна в контексте современной ситуации с экспоненциальным ростом объемов научной информации и информации в целом, поскольку ее решение позволит более эффективно управлять информационными потоками как для отдельно взятого человека, ориентируясь на его информационные потребности и предпочтения, так и для исследовательских групп, лабораторий, институтов, а также на государственном уровне.

В частности, цель настоящей диссертации состоит в разработке модели для информативного реферирования научных текстов на английском языке, которая поможет экономить время исследователям сокращая

объем информации для переработки, и одновременно сохраняя информационную ценность текста.

Задачи данной работы заключаются в:

- Обзор литературы по системам, методам, метрикам качества, данным, и моделям автоматического реферирования текста.
- Подбор данных для проведения экспериментов.
- Оценка наивысшего ROUGE, который можно получить с помощью методов Extractive Text Summarization (ETS).
- Разработка модели автоматического реферирования на основе жадного алгоритма.

Цели и задачи данной диссертации целиком направлены на конечный результат в виде создания модели автоматического реферирования научных текстов на английском языке и состоят из 3 подготовительных (обзор литературы и подбор данных) и непосредственно практических шагов (оценка достижимого качества по метрике ROUGE, и разработка модели). Поэтому я считаю, что выбор целей и задач данной диссертации полностью адекватны.

b) Оригинальность проблемы

Проблема является новой и оригинальной. Насколько мне известно, до публикаций г-на Ахметова и его коллег, идея использования жадного алгоритма для задачи автоматического реферирования текстов не была предложена в литературе.

Отчасти это оправдано тем, что в последнее время фокус исследователя сместился в сторону разработки абстрактных моделей автоматического реферирования текстов. Таким образом, предложение г-на Ахметова не только является новым и оригинальным, но и было представлено очень вовремя, когда современные модели

обеспечения, тестирования результатов и проведения цикла оценки и корректировки.

е) Результаты и обсуждения

В диссертации правильно представлены результаты, полученные в работе, обсуждаются их преимущества, недостатки, ограничения и перспективы. Диссертационный документ грамотно структурирован в представление алгоритма г-на Ахметова, на который было получено авторское свидетельство. В диссертации корректно обсуждены полученные результаты, для чего в каждую главу включен раздел обсуждения.

ф) Заключение и новые выводы

Документ включает раздел "Заключение", в котором перечислены выводы г-на Ахметова, а именно:

- Произведена оценка наилучшего возможного балла ROUGE для методов Автоматического Экстрактивного Реферирования текстов
 - Применены подходы:
 - Переменный поиск соседства (VNS).
 - Подход с использованием жадного алгоритма — для сходимости требуется меньше времени, чем для VNS.
 - VNS инициализированный результатами жадного алгоритма.
 - Полученные результаты:

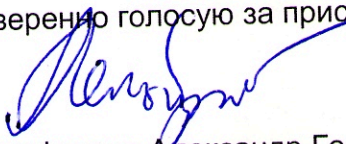


- Наилучшая возможная оценка ROUGE составляет 0,58, а самые современные модели дают всего 0,46.
- Следовательно, есть еще место для развития ETS .
- GreedSum : жадный алгоритм автоматического реферирования текста (ATS)
 - Результаты сопоставимы с результатами современных моделей.
 - Методы ETS все еще имеют потенциал для развития.
 - Ключевым фактором является настройка гиперпараметров.

Поэтому в документе правильно изложены новые выводы и представлено адекватное обсуждение.

2. Заключительные замечания и решение о принятии

Г-н Ахметов опубликовал свои результаты в 2 журнальных статьях и 1 докладе на конференции, что опять же представляет собой независимую внешнюю оценку его теоретических выводов. Исходя из этого, я уверенно голосую за присуждение г-ну Ахметову степени PhD.


Профессор Александр Гельбух, доктор философии.

Член Мексиканской академии наук,
президент Мексиканского общества искусственного интеллекта,
профессор-исследователь и руководитель
лаборатории естественного языка и обработки текста,
Центр вычислительных исследований, Национальный
политехнический институт,