

OVERSEAS RESEARCH SUPERVISOR'S REPORT

on Dissertation Thesis of Oralbekova Dina Orymbayevna on the theme “Development of an automatic speech recognition system based on an end-to-end approach”, submitted for the Ph.D. degree in the specialty 8D06103 – Management information systems

Automatic speech recognition systems are a necessary tool in the management of automated and robotic complex. The development of technology that converts Kazakh speech into text is a big step forward for the economic and social development of the Republic of Kazakhstan

Traditional systems demonstrate fairly high recognition accuracy, but, at the same time, they consist of several independent complex modules, which can cause problems in constructing models. Today, machine learning methods are widely used in the field of speech recognition and, in this regard, the end-to-end method of speech recognition has become widespread. The basic principle of end-to-end models is that they are trained on the basis of big data. From this it is possible to detect the main problem, it concerns the recognition of low-resource languages, which includes the Kazakh language.

Currently, end-to-end systems for recognition of Kazakh speech based on a model using the attention mechanism, which shows promising results, have not been developed. A model with an attention mechanism can work well both with and without language models, while demonstrating a low error in speech recognition. This approach is a promising direction that can be used to develop speech recognition systems with a limited training data.

Thus, it can be assumed that the relevance of the topic of the dissertation work is beyond doubt.

The purpose of the dissertation is to do research and develop a model, method and algorithm to improve the accuracy of recognition of Kazakh speech based on end-to-end architecture.

The scientific novelty of the researcher's work lies in the development of a new speech and text corpus for the Kazakh language, as well as in the implementation of an end-to-end model and method using the attention mechanism for recognition Kazakh speech.

The author made a comparative analysis of modern speech recognition methods based on traditional, hybrid and end-to-end models; developed an extended speech corpus for end-to-end recognition of Kazakh speech; developed an effective method and algorithm of an end-to-end model based on Encoder-Decoder using the attention mechanism to create a recognition system of Kazakh speech; developed an end-to-end architecture and software for recognition of Kazakh speech.

The scientific research conducted by Oralbekova D.O. indicates that the researcher is sufficiently proficient in the methods of scientific analysis, has a technical level of preparedness for conducting complex scientific research, and has extensive knowledge in the field of speech recognition and speech-to-text technology.

On the theme of dissertation 10 scientific articles were published in the international journals indexed in the “Scopus” database as well as during the international conferences, in the magazines and bulletins. Due to the pandemic covid-19 crossed the world, a plan to do the internship program at Universiti Putra Malaysia is cancelled however we still continue our research discussion, implementation and analysis of algorithms, progress presentation and discussion on writing article via video conferencing, zoom.

On the basis of the above mentioned, I consider that the dissertation of D.O. Oralbekova satisfies the requirements for obtaining the degree of Doctor of Philosophy (Ph.D.) in specialty 8D06103 – Management information systems.

Overseas Research Supervisor:
Professor, Dr., Universiti Putra, Malaysia

Mohamed Othman

Штамп:
/Профессор, Д-р. Мохамед Отсман
Кафедра коммуникационных технологий и сетей
Факультет информатики и информационных технологий
Университет Путра Малайзия
43400 УПМ Серданг
Селангор Дарул Эхсан, Малайзия/

ОТЗЫВ ЗАРУБЕЖНОГО НАУЧНОГО КОНСУЛЬТАНТА

на диссертацию Оралбековой Дины Орымбаевны на тему «Разработка системы автоматического распознавания речи на основе интегрального подхода» на соискание степени доктора философии (Ph.D.) по специальности «8D06103 – Management information systems»

Системы автоматического распознавания речи являются необходимым инструментом при управлении кофмплексом автоматизированных и робототехнических устройств. Разработка технологии преобразования казахской речи в текст – большой шаг вперед для экономического и социального развития Республики Казахстан.

Традиционные системы демонстрируют достаточно высокую точность распознавания речи, но при этом они состоят из нескольких независимых сложных модулей, что при построении моделей может вызывать проблемы. На сегодняшний день методы машинного обучения широко используются в области распознавания речи и в связи с этим широкое распространение получил интегральный метод распознавания речи. Основной принцип работы интегральных моделей заключается в том, что они обучаются на основе больших данных. Исходя из этого, можно выделить основную проблему, которая касается распознавания малоресурсных языков, к которым относится и казахский язык.

В настоящее время не разработаны интегральные системы распознавания казахской речи на основе модели с использованием механизма внимания, которая демонстрирует многообещающие результаты. Модель с механизмом внимания может хорошо работать как с языковыми моделями, так и без них, демонстрируя при этом низкую погрешность распознавания речи. Такой подход является перспективным направлением, которое можно использовать для разработки систем распознавания речи при ограниченной обучающей выборки.

Таким образом, можно считать, что актуальность темы диссертации не вызывает сомнений.

Целью диссертации является исследование и разработка модели, метода и алгоритма для повышения точности распознавания казахской речи на основе интегральной архитектуры.

Научная новизна работы исследователя заключается в разработке нового речевого и текстового корпуса для казахского языка, а также в реализации интегральной модели и метода с использованием механизма внимания для распознавания казахской речи.

Автором проведен сравнительный анализ современных методов распознавания речи на основе традиционных, гибридных и интегральных моделей; разработан расширенный речевой корпус для интегрального распознавания казахской речи; разработаны эффективный метод и алгоритм интегральных моделей на базе кодер-декодер с использованием механизма внимания для создания системы распознавания казахской речи; разработаны интегральная архитектура и программное обеспечение для распознавания казахской речи.

Научное исследование, проведенное Оралбековой Д.О., свидетельствует о том, что исследователь в достаточной степени владеет методами научного анализа, имеет

технический уровень подготовленности к проведению сложных научных исследований, обладает обширными знаниями в области распознавания речи и технологии преобразования речи в текст.

По теме диссертации опубликовано 10 научных статей в международных журналах, индексируемых в базе данных «Scopus», а также в ходе международных конференций, в журналах и бюллетенях. Из-за эпидемии COVID-19, охватившей весь мир, планы по прохождению международной программы стажировки в Университете Путра Малайзия были отменены, тем не менее, мы продолжаем наше исследовательское обсуждение, работы по внедрению и анализу алгоритмов, презентацию хода выполнения работ, а также обсуждение написания статьи по видеоконференции и через zoom.

Учитывая вышеизложенное, я считаю, что диссертация Д.О.Оралбековой отвечает требованиям для присвоения ей ученой степени доктора философии (Ph.D.) по специальности 8D06103 – Management information systems

Зарубежный научный консультант:

Профессор, Д-р., Университет Путра Малайзия /подписано/ Мохамед Отсман

Штамп:

/Профессор, Д-р. Мохамед Отсман

Кафедра коммуникационных технологий и сетей

Факультет информатики и информационных технологий

Университет Путра Малайзия

43400 УПМ Серданг

Селангор Дарул Эхсан, Малайзия/

Перевод вышеуказанного документа с английского языка на русский язык выполнен переводчиком Омаровой Маликой Маркисовной, ИИН 930924401943.

Подпись: *Омарова Малика Маркисовна*

Нотариус, свидетельствуя подлинность подписи, не удостоверяет фактов, изложенных в документе, а лишь подтверждает, что подпись сделана определенным лицом.

Республика Казахстан, город Алматы.

Двадцать восьмое апреля две тысячи двадцать второго года.

Я, Бекешбаева Роза Пернебековна, нотариус города Алматы, государственная лицензия № 15022328 выдана Министерством юстиции Республики Казахстан от 25 декабря 2015 года, свидетельствую подлинность подписи переводчика **Омаровой Малики Маркисовны**. Личность, переводчика установлена, дееспособность и полномочия проверены.



Зарегистрировано в реестре за № **4083**
Взыскано: по ставке – 92 тенге
тех. работа – осв. на осн. п.п.1, п.2,
ст.30-1 Закона РК «О нотариате».

Нотариус _____

(Handwritten signature of the notary)



ES6005441220428162415U721509

Нотариаттық іс-әрекеттің бірегей нөмірі / Уникальный номер нотариального действия