

ОТЗЫВ

на автореферат диссертации
Пикалёва Ярослава Сергеевича

на тему *«Совершенствование методов и программных средств
распознавания слитной русской речи»*,
представленной на соискание ученой степени кандидата технических наук
по специальности 05.13.01 – Системный анализ, управление и обработка
информации (по отраслям) (технические науки)

Проблема распознавания слитной речи является одной из самых сложных актуальных задач искусственного интеллекта. Исследования в области распознавания русской речи связаны с рядом трудностей, главными из которых является недостаточное количество аннотированных речевых баз и нормализованных текстов для статистического моделирования языка. Кроме того, различные виды искажений (внешние шумы, вариации произношения, синтаксические отклонения и др.) дополнительно усложняют процесс распознавания.

Диссертация Пикалёва Ярослава Сергеевича, представленная на соискание ученой степени кандидата технических наук, посвящена совершенствованию методов акустического и языкового моделирования, а также распознавания слитной русской речи, что в свете сказанного является **актуальным**.

В диссертации предложены методы распознавания русской речи на основе глубоких нейросетей и проведено исследование их эффективности. Основные результаты диссертации заключаются в следующем:

1. Разработаны методы автоматического построения словаря транскрипций, позволяющие определять позицию ударения, генерировать транскрипцию для слов-исключений и осуществлять практическую транслитерацию с учетом орфоэпических норм русского языка. Для определения позиции ударения модернизирована архитектура нейросети типа Transformer за счет увеличения количества слоёв, использования методов градиентного отсечения и teacher forcing. Генерирование практической транскрипции и транскрипции слов-исключений осуществляется на основе seq2seq модели, которая использует механизм обучения с подкреплением и метод beam-search для выбора наиболее вероятной последовательности символов.

2. Предложены методы получения робастных акустических признаков и обучения акустических моделей на основе глубоких нейронных сетей. Для

параметризации речевого сигнала использовалось объединение ансамбля нейронных сетей с «узким горлом» и архитектура ResNet-50.

3. Разработаны методы классификации для распознавания фонем на основе двух архитектур: сети с временными задержками и двунаправленной сети с долгой кратковременной памятью. При этом в последнем скрытом слое сети применялся механизм внимания.

На основе предложенных методов и моделей построена система автоматического распознавания слитной русской речи и проведена оценка качества ее работы по сравнению с российскими и зарубежными аналогами. Как показали экспериментальные исследования, разработанная система обладает достаточной точностью и превосходит аналоги по скорости распознавания, при этом требуя для своего обучения значительно меньше данных.

Полученные результаты обладают новизной, имеют теоретическое и практическое значение.

Основные результаты опубликованы в издании, входящем в перечень специализированных научных изданий, утвержденных Минобрнауки России (1 публикация)), в изданиях, включенных в систему Российского индекса научного цитирования (2 публикации), и в материалах международных конференций (10 публикаций). Предложенные технологии построения акустических моделей и нормализации текста на основе методов машинного обучения использованы в Институте проблем искусственного интеллекта при выполнении фундаментальной научно-исследовательской работы.

Автореферат дает полное представление о содержании работы. Вместе с тем имеется ряд замечаний:

1) в автореферате диссертации не уделено внимания статистическим языковым моделям.

2) описана модель, используемая для синтаксического анализа, на вход которой подается вектор размерности 64, однако не указано, как формируется этот вектор.

В целом данная диссертация, представленная на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки), является законченной научно-исследовательской работой, имеющей большое как научное, так и практическое значение для разработок энергосберегающих технологий. По своей значимости и научной новизне полученных результатов работа соответствует требованиям, предъявляемым к

кандидатским диссертациям (п. 2.2 «Положения о присуждении ученых степеней»), а ее автор Пикалёв Ярослав Сергеевич заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Доктор технических наук по специальности
05.13.01, 05.13.11

Профессор



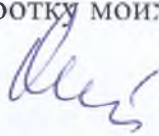
Кузнецов Олег Петрович

Адрес: Россия, 117997, Москва, ул. Профсоюзная, д. 65, ИПУ РАН
тел.: +7 916 458 45 90

E-mail: olpkuz@yandex.ru

ФГБУН Институт проблем управления им. В.А.Трапезникова Российской академии наук, главный научный сотрудник

Согласен на автоматизированную обработку моих персональных данных



Кузнецов Олег Петрович

Подпись ФИО подтверждаю:



Подпись *Кузнецова ОП*
ЗАВЕРЯЮ
ВЕД. ИНЖЕНЕР
ВОВКЕНОВАН