

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)

Актуальность представленной работы обусловлена тем, что распознавание речи (в т.ч. русскоязычной) применяется в широком спектре систем – от голосовых ассистентов до систем «умный дом». При этом в сфере распознавания русскоязычной слитной речи на сегодняшний день успехи достигнуты только в пределах ограниченного словаря в конкретных профессиональных областях. Соответственно, эффективность работы таких систем в условиях разговорной речи оставляет желать лучшего. В связи с этим, задача совершенствования методов и программных средств распознавания слитной русской речи, решению которой посвящена диссертация Пикалёва Я.С., является крайне актуальной.

Объект, предмет, цель и задачи исследования сформулированы логично и четко. Автор достаточно полно освещает комплекс существующих исследований российских и зарубежных специалистов по тематике диссертации. Автореферат диссертации изложен научным стилем, материал исследований структурирован.

Автором создана дикторонезависимая система автоматического распознавания слитной русской речи за счет модернизации алгоритмов системного анализа, обработки и распознавания речевой информации, а также автоматической обработки текстовых данных. **Научная новизна исследования** состоит в том, что автором усовершенствованы нейросетевые методы для решения задач:

- автоматического определения позиции ударения в слове за счет модернизации архитектуры нейросети типа Transformer;
- генерации практических транскрипций англоязычных слов и слов-исключений за счет применения механизма обучения с подкреплением и метода beam-search в модели seq2seq;
- параметризации речевого сигнала за счет объединения ансамбля нейронных сетей с «узким горлом» и архитектуры ResNet-50;

- классификации фонем за счет использования механизма внимания в последнем скрытом слое сети, включающей в себя нейросеть с временными задержками и двунаправленную нейросеть с долгой кратковременной памятью.

Достоверность результатов диссертационного исследования, отраженных в автореферате, подтверждается результатами проведенного численного исследования эффективности функционирования предложенных методов и алгоритмов на независимой статистической выборке, поэтому сомнений не вызывает. Предложенные в диссертации идеи базируются на обобщении передового опыта ведущих российских и зарубежных исследователей в области автоматического распознавания слитной речи. Автором также выполнен анализ качества работы предлагаемой системы автоматического распознавания слитной русской речи по сравнению с российскими и зарубежными аналогами.

Работа прошла широкую апробацию на международных научно-практических конференциях, по результатам выполненного исследования автор имеет достаточное количество публикаций (17 статей) для защиты диссертации на соискание ученой степени кандидата технических наук, в том числе в рецензируемой печати (5 статей).

Научная и практическая значимость работы. Полученные результаты вносят свой вклад в теорию понимания кодирования сигнала на основе глубоких нейронных сетей. Предложенная технология извлечения робастных признаков из речевого сигнала путем его нейросетевой параметризации и метод классификации фонем могут быть использованы при разработке систем голосового управления и поиска по голосовому запросу, а также в системах диктовки с приемлемым уровнем ошибок; разработанные методы построения словаря транскрипций могут найти свое применение при создании систем синтеза речи.

Результаты диссертационной работы использовались в Институте проблем искусственного интеллекта, их практическая значимость подтверждена *справкой о внедрении*.

В качестве **замечания** по работе следует отметить, что в автореферате недостаточно полно описана нейросетевая модель синтаксического анализатора, в частности, остается неясным, что подается ей на вход и что получается на выходе. Кроме того, разработана архитектура сети для классификации ошибок в тексте, что необходимо для процесса нормализации

текста и получения качественной языковой модели. Однако, автор не описал, на какие классы данная модель делит ошибки.

Представленные замечания являются несущественными и не влияют на положительную оценку выполненного диссертационного исследования.

Заключение. Анализ автореферата и публикаций соискателя показывает, что диссертационная работа обладает научной новизной и практической значимостью, соответствует научной специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки). Работа отвечает требованиям п. 2.2 «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям, 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).

Кандидат физико-математических наук по специальности
05.13.18 – Математическое моделирование,
численные методы и комплексы
программ

Юлия Борисовна Линд

Адрес: 450103, г. Уфа, ул. Бехтерева, 3/1
тел.: +7 (917) 803-80-90
E-mail: LindYuB@yandex.ru

ООО «РН-БашНИПИнефть», Ученый секретарь

Согласна на автоматизированную обработку моих персональных данных

Юлия Борисовна Линд

Подпись Ю.Б. Линд удостоверена
Ин. печатью:

