

ОТЗЫВ

официального оппонента на диссертационную работу Кудинова Михаила Сергеевича «Статистическое моделирование русского языка с помощью нейронных сетей», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики»

Актуальность

Диссертационная работа Кудинова М.С. посвящена проблеме статистического моделирования естественного языка. Эффективная статистическая языковая модель необходима во многих практических задачах, связанных с обработкой речи и текста. Примерами таких задач являются автоматическое распознавание речи и рукописного текста, машинный перевод и т.д. Статистическая модель языка должна помогать в предсказании вероятности цепочки слов моделируемого языка и предсказании последующего слова по текущей истории слов. Существует ряд подходов к статистическому моделированию языка. Наиболее популярным подходом долгое время считались сглаженные n -граммные модели языка. В последние годы все большую популярность приобретают языковые модели, основанные на применении технологий машинного обучения и, в частности, глубоких искусственных нейронных сетей. В своей работе Кудинов М.С. исследует применимость современных методов статистического моделирования языка, основанных на глубоких нейронных сетях, к моделированию русского языка. В силу ряда особенностей русского языка многие методы статистического моделирования языка, эффективные для английского и основных мировых языков, оказываются менее эффективными на русскоязычных текстах и требуют существенной модификации. Это свидетельствует об актуальности диссертационной работы Кудинова М.С. по тематике и содержанию.

Научная новизна

В качестве базового метода автор использует языковую модель на основе рекуррентной нейронной сети (RNN). В рамках своего исследования Кудинов М.С. решает две основных задачи.

Первая из них состоит в уменьшении размера словаря языковой модели. Большой размер словаря ведет к росту разреженности данных и низкой производительности вычислений. Для решения этой задачи автором предлагается декомпозиция словаря в словарь начальных форм слов (лемм) и словарь грамматических характеристик словоформ. Последовательное раздельное предсказание начальной формы и грамматической формы позволяет повысить производительность модели и приводит к некоторому уменьшению уровня пословной ошибки распознавания речи (WER).

Вторая проблема, решаемая автором, связана с ограничением возможности рекуррентной нейронной сети для моделирования зависимости между словами последовательности, находящимися далеко в тексте (проблема угасания градиента). Для уменьшения данного эффекта предлагается введение дополнительных признаков, основанных на тематическом разложении документа в рамках вероятностной тематической модели с аддитивной регуляризацией. Подходы, предлагаемые автором, проверяются экспериментально.

В качестве конкретных элементов новизны можно выделить следующие:

1) Предложена статистическая языковая модель с раздельным предсказанием лемм и морфологических признаков словоформ на основе искусственных нейронных сетей с рекуррентной и сверточной архитектурами.

2) Предложено расширение рекуррентной архитектуры за счет признаков, полученных с помощью тематического разложения левостороннего контекста предсказываемого слова.

3) Предложена модель предсказания морфологических признаков слов на основе нейронной сети прямого распространения с одним сверточным слоем.

Результаты, полученные автором, имеют большое значение для прикладных задач, связанных с распознаванием речи и анализом текста.

Содержание работы

Диссертационная работа состоит из введения, 4 глав, заключения и списка литературы. Текст изложен на 106 страницах, в списке литературы 114 наименований. Первая глава носит обзорный характер. В ней дается постановка задачи, анализируются основные метрики качества модели языка и подходы к ее построению.

Вторая глава посвящена подробному анализу языковой модели на основе рекуррентной нейронной сети, проблеме угасания градиента и проблемам применимости данной модели к русскому языку. Данные исследования в значительной степени опираются на современные работы чешского ученого Томаша Миколова и его коллег, посвященные языковому моделированию в системах распознавания речи на основе нейронных сетей. Автором ставится предварительный эксперимент, демонстрирующий эффективность нейросетевой модели языка для предсказания лемм. Применение модели на основе лемм приводит к уменьшению уровня пословной ошибки в эксперименте по переранжированию списка гипотез, возвращаемых внешней системой автоматического распознавания речи.

Третья глава описывает оригинальную гибридную модель языка на основе рекуррентной нейронной сети и вероятностного тематического моделирования. Автором ставится ряд экспериментов с целью выбора оптимального набора гиперпараметров модели: набора аддитивных регуляризаторов, частоты тематического разложения, длины окна анализа и т.д. Эксперимент демонстрирует, что добавление признаков, полученных с помощью тематического моделирования, приводит к снижению перплексии модели (коэффициента неопределенности) на тестовых данных.

В заключительной четвертой главе автором описывается оригинальная модель для предсказания морфологических характеристик слова. Автор использует архитектуру искусственной нейронной сети прямого распространения с одним сверточным слоем. Приводятся результаты эксперимента по ранжированию гипотез, генерируемых системой распознавания речи. Эксперимент показывает, что использование модели для

предсказания морфологических характеристик слов, приводит к снижению уровня пословной ошибки распознавания речи.

Замечания по диссертационной работе

1) В обзорной части диссертации проанализирован достаточно большой объем зарубежной научной литературы, однако, русскоязычные публикации по автоматическому распознаванию русской речи на основе различных моделей языка, следовало бы проанализировать более широко, так в работе совершенно не представлены современные публикации ученых из «Центра речевых технологий», Университета ИТМО, СПИИРАН и других научных организаций РФ.

2) В работе используются некоторые жаргонные слова при переводе ряда англоязычных терминов, которые имеют уже устоявшийся перевод на русский язык, например, вместо “токен” (token) лучше употреблять “слово/словоформа” в соответствующем контексте, вместо “перплексия” (perplexity) – “коэффициент неопределенности”, вместо “отступление” (backoff) – “возврат”, также “семплирование”, “дисконтирование” и т.д.

3) Не совсем ясно какие именно морфологические признаки словоформ использовались в моделях языка, их стоило бы представить в явном виде.

4) Нет описания аудиопризнаков речи и акустической модели, которые применялись в ходе экспериментов по автоматическому распознаванию русской речи при помощи открытой библиотеки Kaldi.

5) Довольно мало обучающих текстовых данных используются при обучении нейросетевых моделей русского языка (всего 2-3 млн слов), для создания более эффективной модели стоит увеличить объем данных.

6) Имеются замечания к оформлению диссертации: в тексте встречаются орфографические и пунктуационные ошибки; на стр. 27 указан рис. 3.1 вместо 1.3; не все публикации в списке литературы оформлены правильно, некоторые не имеют выходных данных (например, [25, 63, 64, 87, 111]); в тексте пропущена ссылка на публикацию [99], в связи с чем в Главе 4 все ссылки с [99] по [115] указаны неверно.

Заключение

Все перечисленные замечания являются несущественными и не влияют на общую положительную оценку диссертационной работы.

Таким образом, работа Кудинова М.С. представляет собой законченную квалификационную работу, в которой исследована проблема статистического моделирования русского языка с помощью нейронных сетей. Основные результаты опубликованы автором в 7 печатных работах, из них 5 – в рецензируемых изданиях из перечня ВАК. Автореферат достаточно полно отражает содержание диссертационной работы.

Считаю, что по уровню актуальности, научной новизны, теоретической и практической значимости диссертационная работа Кудинова М.С. «Статистическое моделирование русского языка с помощью нейронных сетей» в полной мере соответствует требованиям пп. 9–14 Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям ВАК при Минобрнауки РФ, а ее автор Кудинов Михаил Сергеевич заслуживает присуждения ему ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Официальный оппонент:

Карпов Алексей Анатольевич – доктор технических наук, доцент, заведующий лабораторией речевых и многомодальных интерфейсов Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), Санкт-Петербург, 14-я линия, д. 39, e-mail: karpov@iias.spb.su, тел.: (812) 328-0421



Подпись руки Карпов А.А. заверяю,
Начальник отдела кадров СПИИРАН
Д.В.Токарев
20 16 г.