

## Рефераты

УДК 81.322.2

Восприятие искусства, порожденного AI: анализ текстов онлайн-дискуссий. Босоногов С., Суворова А. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 6–23.

В настоящей работе мы анализируем комментарии к трем субреддитам, связанным с искусством, созданным искусственным интеллектом, чтобы понять, как люди воспринимают способность AI создавать искусство, а также темы и настроения дискуссий в контексте широкого использования предобученных моделей. Мы использовали методы вычислительного анализа текстов, такие как тематическое моделирование LDA и анализ тональности с помощью лексиконов тональных слов. В результате мы обнаружили, что дискуссии на технические темы и описания искусства, созданного AI, были в основном позитивными, тогда как дискуссии по социокультурным вопросам были в основном негативными и проходили в субреддите, посвященном защите искусства AI. Результаты показывают, что пользователи Reddit интересуются как художественными, так и социокультурными последствиями искусства, созданного AI, считая его рискованным и сомнительным.

Библ. – 30 назв.

УДК 81.322.2

Вариация в средней валентности слов в русскоязычных книгах в течение ста лет и ее связь с социальными изменениями. Бочкарев В. В., Соловьев В. Д., Нестик Т. А., Шевлякова А. В. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 24–42.

Валентность слов в книгах отражает ситуацию в обществе и позволяет оценить восприятие жизни даже в тех странах и периодах времени, когда прямые исследования благополучия не проводились. Для анализа изменений средней валентности слов в русскоязычных книгах мы используем корпус диахронных текстов Google Books Ngram. Мы показываем, что изменения средней валентности коррелируют с результатами опросов о благосостоянии. Средняя валентность также реагирует на крупные исторические события и социальные изменения. Как и другие подобные исследования, количественные данные

об уровне валентности слов основаны на расчетах с использованием словарей с рейтингами валентности слов. Мы впервые проводим сравнительное исследование на основе ряда наиболее актуальных русских словарей. В результате выяснилось, что полученные результаты зависят от прикладного значения таких словарей и их лексического состава. Это показывает необходимость тщательного выбора словарей для будущих исследований.

Библ. – 46 назв.

УДК 81.322.5

Wav2Vec2 без внимания: нужны ли сети Хопфилда для самообучения представлений речи? Гребенкин Д., Бондаренко И. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семина. ПОМИ, т. 529), СПб., 2023, с. 43–53.

В настоящей работе мы рассматриваем возможность замены многоголового внимания слоями плотной ассоциативной памяти (dense associative memory, DAM) в алгоритме автоматического распознавания речи wav2vec2. Мы рассматриваем гипотезу о том, что концепция современных сетей Хопфилда больше подходит для задачи восстановления недостающих фрагментов звукового сигнала и задачи преобразования речи в текст, чем многоголовое внимание. Наши эксперименты показывают, что модель с новой архитектурой позволяет улучшить качество распознавания речи и может использоваться для предварительного обучения моделей на большом объеме аудиоданных.

Библ. – 27 назв.

УДК 81.322.2

Моноязыковой и межязыковой перенос знаний для задачи тематической классификации. Карпов Д., Бурцев М. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семина. ПОМИ, т. 529), СПб., 2023, с. 54–71.

В этой работе мы исследуем перенос знаний из набора данных RuQTopics. Этот русскоязычный тематический набор данных сочетает в себе большое количество точек данных (361560 текстов с одной меткой, 170930 – с несколькими метками) с обширным охватом классов (76 классов). Мы подготовили этот набор данных на основе необработанных данных Yandex Que. Оценивая модели, обученные на RuQTopics, на шести соответствующих классах из русскоязычного подмножества

набора данных MASSIVE, мы показываем, что набор данных RuQTopics подходит для реальных диалоговых задач, поскольку модели, предназначенные только для русского языка, после обучения на этом наборе данных показывают точность около 85% на этом подмножестве. Мы также обнаружили, что для многоязычного BERT, обученного на RuQTopics и оцененного на тех же шести классах MASSIVE (для всех языков датасета MASSIVE), языковая точность тесно коррелирует (корреляция Спирмена 0.773 с  $p$ -значением  $2.997e-11$ ) с приблизительным размером данных для предобучения BERT на соответствующем языке. В то же время корреляция точности с языковой дистанцией от русского языка не является статистически значимой.

Библ. – 33 назв.

УДК 81.322.2

KRGP: порождение ответов с использованием знаний и персоналии. Косенко Д., Жарикова Д. — В кн.: Исследования по прикладной математике и информатике. П1. (Зап. научн. семина. ПОМИ, т. 529), СПб., 2023, с. 72–85.

Чтобы создать персонализированный ответ, порождающая модель должна учитывать личную информацию о пользователе, заданный вопрос и знания о предметной области. Поэтому необходимо научиться извлекать релевантную информацию, которая поможет порождающей модели составить ответ пользователю. В настоящей работе мы предлагаем разбить процесс на три этапа: отбор релевантных предложений из текстовой базы знаний, отбор наиболее подходящих предложений из текстового описания личности с учетом извлеченных знаний и формирование ответа на основе полученных знаний и персоналии. Мы используем Sentence Transformer и адаптируем алгоритм из работы о модели CLIP для получения контекстозависимых представлений предложений и извлечения наиболее релевантных фрагментов текста из базы знаний. Мы обнаружили, что фокусная функция ошибки (focal loss) показывает лучшие результаты в задачах бинарной классификации персоналии на примере несбалансированного набора данных FoCus. Мы также показали, что text2text Transformer BART хорошо справляется с задачами порождения условного ответа в диалоге. Разработанная нами система заняла первое место в таблице лидеров The 1st Workshop on Customized Chat Grounding Persona.

---

Исходный код разработанных моделей доступен на GitHub по адресу [https://github.com/dmitrymailk/deeppavlov\\_focus](https://github.com/dmitrymailk/deeppavlov_focus)

Библ. – 26 назв.

УДК 81.322.5

Использование модификаций аудиосигнала для подтверждения релевантности акустических признаков. Кочеткова У., Скрелин П., Евдокимова В., Новоселова Д. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 86–101.

В настоящей работе основное внимание уделяется экспериментам с использованием модификаций просодических параметров отдельно и в различных сочетаниях. С этой целью исходные омонимичные нейтральные и иронические высказывания были модифицированы для достижения противоположного эффекта, т.е. для получения иронических высказываний из нейтральных и наоборот. Эти оригинальные омонимичные высказывания зачитывались носителями русского языка в виде мини-текстов и диалогов, подразумевающих нейтральный или иронический смысл, затем извлекались из контекста, устранялись любые маркеры иронического или неиронического значения, затем предлагались носителям языка во время аудио-перцептивных экспериментов. Акустический анализ проводился для пар омонимичных иронических и неиронических высказываний, смысл которых был правильно понят большинством слушателей. Вычислены наиболее яркие акустические особенности иронического смысла. Затем были внесены изменения в пары иронических и неиронических высказываний, содержащих эти особенности. В трех сериях экспериментов временные, динамические и мелодические параметры модифицировались с помощью программного обеспечения Wave Assistant и Praat. Результаты перцептивных экспериментов с модифицированными стимулами не только продемонстрировали преобладание мелодического рисунка, но и выявили комплекс акустических особенностей иронии и указали на трудности модификаций иронических стимулов из-за их спектральной плотности.

Библ. – 23 назв.

УДК 81.322.2

IMAD: мультимодальный диалог, дополненный изображениями. Москворецкий В., Фролов А., Кузнецов Д. — В кн.: Исследования по прикладной математике и информатике. П1. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 102–122.

В настоящее время диалоговые системы достигли отличных результатов при обработке текстовой коммуникации. Однако они еще не могут эффективно дополнять диалог визуальной информацией, что представляет собой серьезную проблему. Более того, существующие модели, включающие изображения в процесс создания диалога, фокусируются на обсуждении самого изображения. Предлагаемый нами подход представляет новый взгляд на мультимодальные диалоговые системы, которые интерпретируют изображение в контексте диалога. Тем самым мы стремимся расширить возможности существующих диалоговых систем и перевести их из одной модальности (текста) в область мультимодальности. Однако для этой задачи не хватает проверенных наборов данных на английском языке, содержащих как изображения, так и контексты диалогов. Таким образом, мы предлагаем двухэтапный подход для автоматического построения набора данных мультимодального диалога. На первом этапе мы используем сходство текста и изображения и сходство предложений, чтобы определить, какие высказывания можно заменить изображением. На втором этапе мы заменяем эти высказывания, выбирая подмножество соответствующих изображений и фильтруя их с помощью модели для ответа на визуальные вопросы (visual question answering). Мы использовали этот подход вместе с дополнительной разметкой для создания набора данных мультимодального диалога IImage Augmented (IMAD), который может служить проверенным набором данных для этой задачи. Кроме того, мы предлагаем базовую модель, обученную на этом наборе данных, которая превосходит модель, обученную на тех же данных без изображений, и BlenderBot.

Библ. – 70 назв.

УДК 81.322.2

Предобучение модели LongT5 для многодокументного реферирования масс-медиа документов на вьетнамском языке. Русначенко Н., Ле Тхе

Ань, Нгуен Нгок Дьеп — В кн.: Исследования по прикладной математике и информатике. II<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 123–139.

Реферирование (summarization) нескольких документов – это задача, направленная на извлечение наиболее важной информации из набора входных документов. Одной из основных проблем в этой задаче является проблема долгосрочных зависимостей. Когда мы имеем дело с текстами, написанными на вьетнамском языке, это также сопровождается специфическим слоговым представлением текста и отсутствием размеченных наборов данных. Недавние достижения в области машинного перевода привели к значительному росту использования архитектуры, известной как *Transformer*. После предварительного обучения на больших объемах необработанных текстов трансформеры позволяют получить глубокие знания о текстах на естественном языке. В настоящей работе мы рассматриваем результаты применения языковых моделей для решения задач реферирования текста, включая важные модели реферирования текстов на вьетнамском языке. Мы используем модель LongT5 для предварительного обучения, а затем дообучаем её для задачи реферирования текста из нескольких документов на вьетнамском языке. Мы анализируем полученную модель и экспериментируем с многодокументными наборами данных на вьетнамском языке, включая ViMs, VMDS и VLSP2022. В результате мы приходим к выводу о том, что использование модели на основе Transformer, предварительно обученной на большом количестве неразмеченных вьетнамских текстов, позволяет нам достичь многообещающих результатов с дальнейшим улучшением за счет дообучения в пределах небольшого количества реферированных вручную текстов. Предварительно обученная модель, используемая в разделе экспериментов, доступна онлайн по адресу <https://github.com/nicolay-r/ViLongT5>.

Библ. – 33 назв.

УДК 81.322.2

Формулы читабельности для трех уровней российских школьных учебников. Соловьев В., Иванов В., Солнышкина М. — В кн.: Исследования по прикладной математике и информатике. II<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 140–156.

В настоящей работе мы предлагаем новую формулу сложности текста, направленную на оценку сложности русских школьных учебников.

В качестве данных для обучения и валидации мы использовали аннотированный Российский академический корпус, содержащий более 5 миллионов токенов, а также использовали в исследовании методы машинного обучения. Значения четырех параметров в каждом из 154 текстов, использованных для исследования, были измерены с помощью инструментов библиотеки Spacy. Сравнительный анализ новой и существующих формул сложности показывает, что различия между ними показательны, и новые формулы дают более точные результаты. Это исследование расширяет наше понимание взаимозависимости между частотой слов и сложностью текста и обеспечивает основу для эффективного применения лексических частотных моделей в исследованиях сложности дискурса. Результаты могут быть использованы авторами учебников и разработчиками тестов для выбора и модификации текстов для определенных категорий читателей. Другие области применения включают дизайн веб-сайтов, опросы и семантический анализ социальных сетей.

Библ. – 45 назв.

#### УДК 81.322.2

Что модели для порождения изображений знают о языках мира? Фирсанова В. — В кн.: Исследования по прикладной математике и информатике. II. (Зап. научн. семина. ПОМИ, т. 529), СПб., 2023, с. 157–175.

Модели, порождающие изображения по тексту (text-to-image), используют заданные пользователем запросы для создания изображений. Такие text-to-image модели, как DALL-E 2, Imagen, Stable Diffusion и Midjourney, могут генерировать фотореалистичные или похожие на нарисованные человеком изображения. Помимо имитации человеческого искусства, большие text-to-image модели научились создавать комбинации пикселей, напоминающие подписи на естественных языках. Например, сгенерированное изображение может содержать фигуру животного и комбинацию символов, напоминающую нам слова на естественном языке, описывающие биологическое название этого вида. Хотя слова, иногда появляющиеся на сгенерированных изображениях, могут быть читабельными для человека, они не укоренены в словарях естественного языка и не имеют смысла для лингвистов. В то же

время мы обнаруживаем, что семиотический и лингвистический анализ так называемого скрытого словаря моделей преобразования текста в изображение внесет вклад в области объяснимого искусственного интеллекта и инженерии запросов. Результаты этого анализа можно использовать для снижения рисков применения таких моделей при решении реальных задач и для обнаружения дипфейков. Предлагаемое исследование является одной из первых попыток анализа text-to-image моделей с точки зрения семиотики и лингвистики. Наш подход предполагает инженерии запросов, создание изображений и сравнительный анализ. Исходный код, сгенерированные изображения и запросы доступны по адресу

<https://github.com/vifirsanova/text-to-image-explainable>

Библ. – 34 назв.

УДК 004.852

Объединение предсказаний улучшает понимание мультимодального рекламного контента. Алексеев А., Савченко А., Тутубалина Е., Мясников Е., Николенко С. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семина. ПОМИ, т. 529), СПб., 2023, с. 176–196.

Рекламная индустрия использует несколько модальностей контента для доставки подразумеваемых сообщений: изображения, видео, текст, музыка, и все это вместе взятое. “Декодирование” сообщения, подразумеваемого в мультимодальном контенте, часто требует анализа как текстовых, так и визуальных его компонентов. Мы изучаем задачи распознавания мультимодального символизма, обнаружения тем и классификации тональности. Руководствуясь разницей в частях сообщения, передаваемых разными модальностями в рекламных объявлениях, мы обучаем отдельные модели для изображений и текстов и значительно улучшаем результаты, смешивая прогнозы на основе изображений и текста (с автоматически распознанным текстом), что дает экспериментальное подтверждение нашего подхода.

Библ. – 69 назв.



УДК 519.83

Аксиомы рациональности фон Неймана–Моргенштерна и неравенства в анализе. Осипов Н. Н. — В кн.: Исследования по прикладной математике и информатике. П<sub>1</sub>. (Зап. научн. семин. ПОМИ, т. 529), СПб., 2023, с. 197–217.

Меняющееся благосостояние агента, делающего ставки на результаты бросков честной монеты, является классическим примером случайного процесса с мартингальным свойством. В такой игре не существует стратегии, которая бы давала положительное математическое ожидание прибыли. Однако задача о том, можно ли рационально выбрать стратегию, отличную от бездействия, остается осмысленной и нетривиальной: оказывается, что есть некоторый “зазор” между полным отказом от игры и полностью *нерациональным* экономическим поведением, которое нарушает базовые аксиомы рациональности фон Неймана–Моргенштерна. Решая задачу об описании этого “зазора” и поиске в рамках него оптимальных стратегий, мы приходим к функциям Беллмана, которые ранее возникали в решении полностью абстрактных задач о поиске точных констант в неравенствах из анализа. Тем самым мы получим естественную экономическую интерпретацию для этих неравенств и связанных с ними функций Беллмана.

Библ. – 19 назв.