

ГЕНЕРАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СЕМАНТИЧЕСКИ НАГРУЖЕННЫХ БОЛЬШИХ ХУДОЖЕСТВЕННЫХ ТЕКСТОВ

Вологина Е.А. (Пятигорский государственный университет)

Научный руководитель – кандидат экономических наук, доцент кафедры информационно-коммуникационных технологий, математики и информационной безопасности Тимченко О.В. (Пятигорский государственный университет)

Введение. Человеческий мыслительный процесс так же, как и язык, служащий инструментом для передачи мысли, очень гибок и с трудом поддается формализации. По этой причине семантический анализ справедливо расценивается как наиболее сложный в контексте автоматической обработки текста. Целью данной работы является рассмотрение проблем извлечения семантической нагрузки из больших объемов текста, в частности, текстов сказок, фэнтези и текстов, изобилующих средствами выразительности, проведение сравнительного анализа существующих нейросетевых технологий генерации изображений.

Основная часть. На данный момент существует огромное количество различных нейронных сетей для генерации изображений на основе введенного текста. Однако до сих пор существует проблема полного и осмысленного понимания больших объемов текста, что необходимо для генерации изображений на его основе. Эмпирическим методом был проведен анализ наиболее популярных генераторов изображений по текстовым данным, результаты исследования показали, что в настоящее время достичь желаемого качества генерации изображений по расширенному текстовому описанию невозможно. Причиной этого служит то, что нейронные сети неполно учитывают семантическую структуру текста. Это может быть связано с тем, что нейросети способны обнаруживать различные латентные связи между словами в тексте, но не способны дать объяснение для таких связей. [1] Следовательно, методы, которые в основном опираются на обработку большого корпуса текстов, недостаточны для глубинного понимания смысловой нагрузки художественного текста. Для глубинного понимания смысла текста требуется менее формальное представление знаний о мире и взаимосвязях в нем. [1]

Таким образом, необходима доработка нейронных сетей до такого уровня, когда нейросеть сможет находить формализованное представление смысла текста. Одним из предложений в решении задачи извлечения смысловой нагрузки из больших текстовых данных является синтез наиболее эффективных технологий в области генерации изображений на основе больших текстов, использующих архитектуру Transformer. Языковая модель Transformer весьма эффективна для извлечения семантической нагрузки даже из больших текстов, однако существующие генераторы изображений обучены на массивах изображений и коротких подписях к ним, причем в подписи не включены художественные средства выразительности (метафоры, аллегории). Для улучшения качества генерации необходим датасет, состоящий из сложных языковых конструкций и иллюстраций к ним, а также необходимо разработать алгоритм, который бы учитывал не только семантику исходного текста, но и композиционную структуру самого изображения. [2] После обучения возможно достижение уровня художников-иллюстраторов. Данная работа посвящена разработке архитектуры приложения генерации иллюстраций семантически нагруженных текстов на примере народных сказок на основе моделей трансформеров, кроме того рассматривается специфика обучающего набора данных для решения задачи генерации.

Архитектура системы: последовательная синергия нейронных сетей. Первая нейронная сеть работает по такому принципу: на вход поступает первичный короткий текст, с возможностью выбора сюжета, имен героев и места действия, для дальнейшей генерации сказки. Следующим этапом является извлечение аннотации из сгенерированного текста. Так как особенностью является то, что текст включает в себя

множество средств выразительности, необходимо чтобы нейросеть учитывала это при сборе аннотации, что и является следующей задачей в разработке данной нейронной сети. Конечной целью является генерация иллюстраций полностью учитывающей все особенности текста. На данный момент идет работа с датасетом, таким образом, чтобы сгенерированные тексты соответствовали стилистике народных сказок и, следовательно, содержали специфическую лексику и средства выразительности.

Выводы. Исходя из приведенных выше тезисов, можно сделать вывод, что к настоящему времени генеративные нейронные сети пока что не в состоянии корректно обрабатывать и воспринимать большие специфические тексты. Одним из вариантов решений видится не только создание специализированного датасета, но и синергия нейросетей генерации текста и изображения.

Список использованных источников:

1. Батура Т.В. Семантический анализ и способы представления смысла текста в компьютерной лингвистике // Программные продукты и системы. 2016. №4 (116).
2. Петрин Д.А., Белов Ю.С. Улучшение качества моделей машинного обучения в задачах классификации изображений на основе подходов извлечения признаков и точной настройки модели // Электронный журнал: наука, техника и образование. 2020. № 1 (28). С. 104- 111.