

Статья: How Contextual are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings.

Автор исследования: Артем Стрельцов, группа 182

- Год выпуска статьи 2019. Представлена на конференции EMNLP 2019 в виде презентации с выступлением.
- Автор -- из университета Стэнфорда, PhD. Работает в Фейсбуке. Специализируется на representation learning и оценке моделей.

Другие работы:

- **On the Opportunities and Risks of Foundation Models.**
- **Dynaboard: An Evaluation-As-A-Service Platform for Holistic Next-Generation Benchmarking**
- **Attention Flows are Shapley Value Explanations.**
- **Utility is in the Eye of the User: A Critique of NLP Leaderboards.**
- **Is Your Classifier Actually Biased? Measuring Fairness under Uncertainty with Bernstein Bounds.**
- **How Contextual are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings.**
- **Rotate *King* to get *Queen*: Word Relationships as Orthogonal Transformations in Embedding Space.**
- **Understanding Undesirable Word Embedding Associations.**
- **Towards Understanding Linear Word Analogies.**
- **Unsupervised Random Walk Sentence Embeddings: A Strong but Simple Baseline.**

Я бы не сказал, что это прямо случайная находка. Автор давно занимается исследованием NLP-моделей, и в целом достаточно логично сравнить статические и контекстуальные эмбединги.

- Базовыми можно, очевидно выделить статьи, посвященные BERT (Devlin et al., 2018), ELMo (Peters et al., 2018) и GPT-2 (Alec Radford et al., 2019), поскольку для исследований преимущественно взяты модели из этих работ.

Вдохновением для автора, как мне кажется, являются работы Ian Tenney et al. (What do you learn from context? probing for sentence structure in contextualized word representations, 2019) и David Mimno and Laure Thompson (The strange geometry of skip-gram with negative sampling, 2017). Первая статья концептуально похожа на ту, которую мы рассматриваем, однако автор сам замечает, что там акцент делается лишь на семантическую и синтаксическую информацию в представлениях, и утверждает, что больше его статья идейно похожа на исследование по геометрии статических эмбедингов (то есть на вторую статью).

- У статьи около 169 цитирований. В частности много работ, которые связаны с исследованием геометрии представлений и их контекстуальности:
 - [Circles are like Ellipses, or Ellipses are like Circles? Measuring the Degree of Asymmetry](#)

[of Static and Contextual Embeddings and the Implications to Representation Learning](#)

- [Effects of Architecture and Training on Embedding Geometry and Feature Discriminability in BERT](#)
- [Assessing Lexical-Semantic Regularities in Portuguese Word Embeddings](#)
- [A comprehensive comparative evaluation and analysis of Distributional Semantic](#)
- [A Comparative Study on Structural and Semantic Properties of Sentence Embeddings](#)

И, например, во второй статье, например, данная работа упоминается как одна из ближайших по теме.

- У данной статьи не то чтобы есть какие-либо прямые конкуренты. Вероятно, это связано с тем, что статья скорее предоставляет сравнительный обзор и исследование контекстуальности представлений, а также предлагает метод оценки. Можно заметить, к примеру, что какие-то закономерности среди контекстуальных моделей в плане косинусной близости на различных были предложены и ранее, например, в статье ["Dissecting Contextual Word Embeddings: Architecture and Representation"](#).
- В качестве продолжения статьи можно было бы предложить введение штрафа за анизотропию (в целом авторы это упоминают). Можно также попробовать использовать это как некоторое подобие "регуляризации" для обучения более или менее контекст-специфичных эмбедингов, например, увеличивать этот штраф, если нужно, чтобы эмбединги были именно task-specific. Возможно, это даже не очень сложно реализуется, поскольку автор уже предложил некоторый способ оценки анизотропии
- Тут сложно сказать о какой-то применимости статьи в продакшне. Так как исследование больше сравнительное, то оно несет ценность скорее вспомогательную, несколько углубляя понимание того, как ведут себя модели для эмбедингов на различных слоях. Несомненно, это может принести некоторый результат при использовании указанных в статье моделей, ведь понимание того, как все устроено и как работает, -- это одна из самых важнейших составляющих.