УДК 004.852

ОБУЧЕНИЕ УПОРЯДОЧЕННЫХ ПРЕДСТАВЛЕНИЙ БЕЗ РАЗМЕТКИ

Рыков А.Г. (ИТМО)
Научный руководитель – кандидат технических наук, доцент Муравьев С.Б.
(ИТМО)

Введение. Задача обучения представлений играет ключевую роль в машинном обучении, поскольку неструктурированные данные, такие как изображения или текст, необходимо преобразовать в векторный вид для дальнейшего обучения моделей. Нейронные сети чаще всего используются для получения таких векторных представлений. Сохранение структурных свойств данных, таких как иерархия или порядок, в пространстве представлений позволяет создавать более эффективные векторы, которые улучшают качество работы модели на этапе её применения [1]. Однако в настоящее время практически отсутствуют методы обучения упорядоченных представлений с использованием глубокого обучения в условиях отсутствия размеченных данных.

Основная часть. В качестве альтернативы существующим методам, которые полагаются на явные сигналы для обучения с учителем [1], предлагается использовать подходы, основанные на самообучении (self-supervised learning). Многие из этих методов можно интерпретировать как задачу аппроксимации функции семантической схожести между объектами, где информация о схожести представлена в виде неполной матрицы смежности [2].

Эта задача тесно связана с теорией спектра графов, которая активно применяется в нелинейных методах снижения размерности и моделирования многообразий. В рамках этой теории используется собственное разложение матрицы смежности, которая обычно задаётся некоторой функцией ядра. Собственные векторы (или функции) могут быть упорядочены в соответствии с их собственными значениями, что позволяет отразить структурные свойства графа, такие как связность и иерархия узлов [3].

Для адаптации идей спектра графов уже разработаны методы, позволяющие аппроксимировать собственные функции ядра с помощью нейронных сетей [4][5]. Аппроксимация собственных функций оператора семантической схожести, в качестве которого может выступать ядро аугментации (в задачах самообучения) или более простые ядра в исходном пространстве, позволяет получить упорядоченные представления данных, которые сохраняют свойства выбранного ядра.

Выводы. Предложенный метод позволяет обучать представления с сохранением упорядоченности информации без необходимости использования явных сигналов. Такие представления могут быть полезны для последующего анализа, например, для задач кластеризации или построения иерархической разметки.

Список использованных источников:

1. Kusupati A. et al. Matryoshka representation learning //Advances in Neural Information Processing Systems. – 2022. – Т. 35. – С. 30233-30249.
2. Munkhoeva M., Oseledets I. Neural harmonics: Bridging spectral embedding and matrix completion in self-supervised learning //Advances in Neural Information Processing Systems. –

2023. – T. 36. – C. 60712-60723.

3. Von Luxburg U. A tutorial on spectral clustering //Statistics and computing. – 2007. – T. 17. – C. 395-416.

4. Deng Z., Shi J., Zhu J. Neuralef: Deconstructing kernels by deep neural networks //International Conference on Machine Learning. – PMLR, 2022. – C. 4976-4992.

5. Deng Z. et al. Neural eigenfunctions are structured representation learners //arXiv preprint arXiv:2210.12637. – 2022.