

Методы снижения затрачиваемых ресурсов для вычисления персистентных гомологий на медицинских данных

Панкова К.С. (ИТМО), Добренко Д.А. (ИТМО)

Научный руководитель - Кандидат технических наук, доцент Гусарова Н.Ф. (ИТМО)

Введение. Персистентная гомология (ПГ) — это метод, используемый в топологическом анализе данных для изучения качественных признаков данных, которые сохраняются в различных масштабах. Метод устойчив к искажениям входных данных, не зависит от размеров и координат, обеспечивает компактное представление признаков входных данных. Область вычислений ПГ быстро развивается, новые алгоритмы и реализации программного обеспечения обновляются и выпускаются быстрыми темпами. [1]

Основная часть. Топологический анализ данных и, в частности, персистентные гомологии (ПГ) являются перспективным инструментом мониторинга информационных процессов на входе и внутри нейронной сети (НС). Вычисление ПГ для датасетов больших размеров является очень дорогой вычислительной операцией. [1] Для снижения вычислительной сложности оценки ПГ используются аппроксимирующие методы двух типов:

- 1) кластерная редукция, например, основанная на методе k-means [2]
- 2) редукция через симплициальный коллапс комплекса, используемая в методе SimBa [3].

В работе проведен теоретический анализ существующих методов вычисления ПГ. Проведены экспериментальные оценки времени и вычислительной сложности расчета ПГ для различных случаев на медицинских данных.

Выводы. В работе проведен анализ существующих методов снижения затрачиваемых ресурсов для вычисления персистентных гомологий и сравнение результатов их работы на нескольких медицинских датасетах.

Список использованных источников:

1. Otter, N., Porter, M.A., Tillmann, U. et al. A roadmap for the computation of persistent homology. EPJ Data Sci. 6, 17 (2017). // С 2-3
2. A. Moitra, N. Malott, and P. A. Wilsey, “Cluster-based data reduction for persistent homology,” 2018 IEEE International Conference on Big Data, ser. Big Data 2018, Dec. 2018 // С. 327–334.
3. T. K. Dey, D. Shi, and Y. Wang, “Simba: An efficient tool for approximating rips-filtration persistence via simplicial batchcollapse,” 24th Annual European Symposium on Algorithms (ESA 2016), 2016