

Рецензия

на выпускную квалификационную работу Абрамовой Анастасии Николаевны по теме „Метод Монте-Карло по схеме марковской цепи для оценки вероятности редких событий в задачах биоинформатики”

Определение наиболее вероятных пептидов на основе данных, полученных посредством масс-спектрометрии — важная задача современной протеомики. Существующие алгоритмы в этой области имеют немало ограничений, и представленная на рецензию работа делает шаг в сторону большей универсальности.

Представленный в теоретической части алгоритм является существенным улучшением алгоритма MS-DPR. Добавлена оценка дисперсии, не требующая больших вычислительных ресурсов и позволившая ввести обоснованный критерий останова. Также применен алгоритм Ванга-Ландау для оценки весов вместо итеративной их оценки. Помимо этого, проделана большая работа по теоретическому обоснованию непрерывного случая, что актуально ввиду высокой точности современных инструментов. Отдельного внимания заслуживает качественная работа с литературой: приведен обзор опубликованных в последнее десятилетие статей на тему оценки дисперсии в марковских цепях.

В практической части сделана попытка проверить корректность реализации путем сравнения результатов с оценками по методу Монте-Карло. Наглядно показано преимущество метода существенной выборки. Предъявлен успешный практический эксперимент с данными GNPS, показавший увеличение числа идентификаций при одинаковом уровне FDR. Однако хотелось бы, чтобы автор самостоятельно реализовала алгоритмы MS-GF и/или MS-DPR, а не полагалась на программное обеспечение, значительно изменившееся с момента выхода соответствующих статей. Также упущены некоторые детали, такие как выбранная scoring функция и описание построения теоретических спектров.

Несмотря на замечания относительно практической части считаю, что работа заслуживает оценки „отлично”.

Рецензент Тарасов А. Л.