

УДК 004.85

ПРЕДСКАЗАНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ БЕЛКА ПРИ ПОМОЩИ ГЛУБОКОГО ОБУЧЕНИЯ

Артемьев А.Д. (Университет ИТМО)

Научный руководитель – кандидат технических наук, доцент Маркина Т.А.
(Университет ИТМО)

Введение. Предсказание пространственной структуры белка является одной из важнейших [1] проблем в биологии и одновременно одной из сложнейших из-за отсутствия аналитического решения. Существующие подходы используют оптимизационные [2] алгоритмы, однако проблемой является вычислительная сложность. Одним из способов предсказать структуру белка по первичной последовательности аминокислот — использование нейронных сетей и алгоритмов обучения с подкреплением.

Основная часть. Для решения задачи предлагается рассмотрение следующих моментов:

- биохимические аспекты белков для корректного описания элементов и связей,
- физические свойства, описывающие математическую модель взаимодействия аминокислот, позволяющие определить количественные свойства, такие как энергия связей; в данном случае такая модель учитывает гидрофобные и гидрофильные аминокислотные остатки на двухмерной сетке,
- выбор и архитектура модели машинного обучения, позволяющей учесть все закономерности, включая дизайн нейросетевого агента и среды,
- выбор алгоритма обучения с подкреплением для обучения нейросетевого агента для более оптимального взаимодействия со средой,
- реализация библиотеки в виде программного кода для проведения экспериментов и верификации результатов на известных бенчмарках.

Таким образом, предполагается создание алгоритма, который для заданной последовательности аминокислот будет настраивать параметры нейронной сети агента. Агент, в свою очередь, за каждые действия по расстановке узлов в пространстве получает награду, зависящую от энергии взаимодействия аминокислот. После обучения можно применять данного обученного агента для новых последовательностей, что позволит быстрее и точнее получать пространственное устройство белков, чем это происходит сейчас при помощи вычислительных методов и в лаборатории.

Выводы. Результаты реализации описанного подхода для простейшей математической модели белка позволяют говорить о значительном улучшении качества предсказания структуры относительно известных оптимизационных алгоритмов, что говорит о необходимости продолжения исследования проблемы с использованием более сложных моделей протеинов и нейронных сетей.

Список использованных источников:

1. Абдуллаева Н.Ф. Влияние болезни Паркинсона на фолдинг белка. Инновационные научные исследования, 2022. 2-2(16), 11–17.
2. Бекмуратов Т.Ф. Реализация муравьиного алгоритма фолдинга белков методами программных агентов в распределенных системах / Т. Ф. Бекмуратов, Р. К. Базаров, Д. К. Базаров // Проблемы вычислительной и прикладной математики, 2017. – No 2(8). – С. 103-113. – EDN YPASKR.