Qeios

Peer Review

Review of: "Protein Structure Prediction in the 3D HP Model Using Deep Reinforcement Learning"

Meghana Palukuri[1]

1. University of Texas at Austin, United States

Overall, the paper is well-written with a good description of the methods, which provide a good refresher on RL basics, making it suitable for audiences newer to RL as well. The paper can benefit from strengthening the related work section by providing descriptions of relevant work to highlight this paper's contributions further, and the results and discussion sections can be improved by providing and referencing more supporting evidence, such as tables and figures. Please find below some specific suggestions to help improve the paper:

- The abstract and introduction can benefit from a brief summary of the results in quantitative terms.
- While using references in the first person, please use XYZ et al. instead of just using the reference number, e.g., page 2 - section 2.1, in section 2.2.2 by [16]
- Please explain T in equation 4.
- A suggestion on the method - Why not provide rewards at each individual step instead of at the end to potentially guide the amino acid placement process better?
- Apart from computational efficiency and empirical results, can you provide more theoretical explanation on how the reservoir-based architecture is suitable for this problem - e.g., by expanding on how it captures temporal dependencies in the protein folding process that is unique to this architecture.
- There are some typos in the paper - No space between the words 'dimension' and 'using' in the LSTM section.
- Figure 3 states LSTM-A without prior mention of this acronym in the methods section text.
- Figure 3 legend states that there are 8 heads for the attention layer, but the methods description states 4. Please clarify.

doi.org/10.32388/H35GE7

- Please expand acronyms - BKV, OLH in the results section text.
- In the first paragraph of the results section, many statements are made about the results without citing the table or figure with the numerical results, e.g., the statement - 'This model was effective for shorter sequences, demonstrating reasonable performance in achieving optimal energy states.' Another example in the discussion section - 'When compared to core-directed chain growth methods and hybrid evolutionary algorithms, both LSTM-A and FFNN-R show comparable or superior performance in terms of final energy values while requiring less parameter tuning and fewer computational resources.'
- Please add a description for Table 3 in the results section and cite the table.
- It would be helpful to compute an aggregate score of the energy values across all the sequences experimented with so that comparison across algorithms is easy.
- The graph titles and axis text are too small in Figure 7.
- Please indicate in the Figure 7 legend which graph corresponds to which head for better readability.
- The method BILS is mentioned in the results without prior introduction. Please provide a description for each of the methods used for comparison in the related work section.
- The beginning of the second-to-last paragraph in the discussion section on limitations can be moved to the limitations section and summarized to reduce redundancy.
- What were the training, validation, and test datasets used in the experiments? Please elaborate.
- Please provide a brief discussion on the figures and tables provided in the appendix.
- The GitHub repo mentions that the code is built on the work from the following paper. While there is a mention of this paper in section 2.1, please provide a brief description of this paper and how the current work enhances this paper: Yang, K., Huang, H., Vandans, O., Murali, A., Tian, F., Yap, R. H., & Dai, L. (2022). Applying deep reinforcement learning to the HP model for protein structure prediction. Physica a Statistical Mechanics and Its Applications, 609, 128395. https://doi.org/10.1016/j.physa.2022.128395

Declarations

Potential competing interests: No potential competing interests to declare.