

Review of: "ProteinGLUE: A multi-task benchmark suite for self-supervised protein modeling."

Jianfeng Gao¹

¹ Wuhan University

Potential competing interests: The author(s) declared that no potential competing interests exist.

In the manuscript entitled "ProteinGLUE: A multi-task benchmark suite for self-supervised protein modeling", the authors introduce the ProteinGLUE benchmark for the evaluation of protein representations: a set of seven tasks for evaluating learned protein representations. The authors also offer reference code, and they provide two baseline models with hyperparameters specifically trained for these benchmarks. Pre-training was done on two tasks, masked symbol prediction and next sentence prediction. The authors show that pre-training yields higher performance on a variety of downstream tasks such as secondary structure and protein interaction interface prediction, compared to no pre-training. As self-supervised language modeling is getting more attention, the work sounds interesting.

However, the paper need to be revised carefully further for the following reasons. Firstly, the reliability, universality and stability of ProteinGLUE Benchmark Set need further confirmation. Secondly, the English must be improved in terms of grammar, formatting, sentence structure, Italicization, referencing and citation format. Together, I suggest that this paper is able to be accepted for publication with Minor Revision.