EnzBert: Deep attention network for enzyme class predictions
Nicolas Buton, Yann Le Cunff, François Coste

To cite this version:
Nicolas Buton, Yann Le Cunff, François Coste. EnzBert: Deep attention network for enzyme class predictions. JOBIM 2022 - Journées Ouvertes en Biologie, Informatique et Mathématiques, Jul 2022, Rennes, France. pp.1-1. hal-03780557

HAL Id: hal-03780557
https://hal.science/hal-03780557
Submitted on 19 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
EnzBert: Deep attention network for enzyme class predictions

1. Task
1.1 Functional annotation of protein sequences

2. Methods
2.1 Attention mechanism and Transformer
Attention is the basis of Transformers, for each position it gives weights to other residues.

2.2 Two phases training
Unsupervised pre-training on all proteins and supervised fine-tuning step on enzyme classification

2.3 Interpretability method
Attention maps from multiple layers and multiple heads. A new interpretability method to get residues’ importance scores: attention aggregation.

3. Results
3.1 Enzyme classification
Comparison with state-of-the-art method at level 1 and 2 of the enzyme commission (EC) number on ECPred40 dataset.

3.2 Interpretability
Benchmark: Evaluated on the identification of catalytic residues on enzymes with respect to different interpretability methods.

4. Conclusion
- State-of-the-art prediction on enzymes’ classes from sequences only for our model EnzBert
- New interpretability method for Transformers that works very well on enzymes
- Prospects: Considering the hierarchy of Enzyme Commission (EC) number may improve our results

Contact
email: nicolas.buton@irisa.fr

Bibliography